



Evolution de l'agrégation des galaxies dans le sondage VIMOS-VLT Deep Survey

Baptiste Meneux

► To cite this version:

Baptiste Meneux. Evolution de l'agrégation des galaxies dans le sondage VIMOS-VLT Deep Survey. Astrophysique [astro-ph]. Université de Provence - Aix-Marseille I, 2005. Français. NNT : . tel-00011320

HAL Id: tel-00011320

<https://theses.hal.science/tel-00011320>

Submitted on 8 Jan 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DE PROVENCE (AIX-MARSEILLE I)
ECOLE DOCTORALE PHYSIQUE ET SCIENCES DE LA MATIÈRE
Laboratoire d'Astrophysique de Marseille

THÈSE

présentée pour obtenir le grade de

Docteur de l'Université de Provence (Aix-Marseille I)
Spécialité *Rayonnement, Plasma et Astrophysique*

par

Baptiste MENEUX

ÉVOLUTION DE L'AGRÉGATION DES GALAXIES DANS LE SONDAGE VIMOS-VLT *Deep Survey*

*Soutenue publiquement le vendredi 28 octobre 2005
Au Laboratoire d'Astrophysique de Marseille
Devant le Jury composé de*

Philippe AMRAM	Président
Françoise COMBES	Rapporteuse
François R. BOUCHET	Rapporteur
Sophie MAUROGORDATO	Examinatrice
Luigi GUZZO	Examineur
Olivier LE FÈVRE	Directeur

Remerciements

Un nouveau cycle de trois ans ce termine. Après le lycée Jean Monnet à Cognac, mes classes préparatoires au lycée Camille Guérin de Poitiers, et l'Ecole Nationale Supérieure de Physique de Marseille, voici maintenant trois années de thèse qui s'achèvent dans cette belle cité phocéenne !! Je ne suis jamais très bon dans les remerciements mais je tiens ici à citer les nombreuses personnes sans qui ces trois dernières années ne se seraient pas aussi bien passées.

Tout d'abord, je remercie Olivier Le Fèvre qui m'a proposé ce très beau sujet de thèse. Quoi de mieux que de décrire les structures dans l'Univers et suivre leur évolution jusqu'à une époque où celui-ci n'était âgé que de $\sim 30\%$ de son âge actuel, quand depuis tout petit on plonge ses yeux vers le ciel à se demander comment on en est arrivé là. Les choses du quotidien prennent alors une toute autre dimension ...

J'ai tout particulièrement apprécié la confiance qu'Olivier m'a accordée durant ces trois années (et même un peu plus avec le DEA) ainsi que ses encouragements dans les moments difficiles. Je le remercie aussi pour la relecture de cette thèse et ses conseils qui m'ont permis d'aboutir aux 200 et quelques pages que vous, lecteurs, tenez maintenant entre vos mains.

Et puis, trois ans après, je le remercie encore de m'avoir permis d'aller passer 4 nuits d'observation à Hawaii, au CFHT, à 4200m d'altitude (avec "la moitié de l'atmosphère sous mes pieds" dicit Olivier dans un de ces mails en mai 2002), lors de mon stage de DEA. Là où le rêve de gosse a un peu commencé à devenir réalité ...

Je remercie Luigi Guzzo pour son aide, son soutien et le très bon accueil que j'ai eu à chacun de mes séjours à Merate, petite bourgade entre Milan et le magnifique lac de Côme. L'observatoire est très confortable et ce sera avec plaisir que j'y logerai à partir de janvier 2006 et jusqu'à l'été.

Je remercie Françoise Combes, première femme astronome à être élue à l'Académie des Sciences, et François R. Bouchet, responsable du programme Planck en France, d'avoir accepté d'être mes rapporteurs. Je remercie Sophie Maurogordato et Luigi Guzzo pour leur rôle d'examineur dans mon jury, et enfin Philippe Amram pour l'avoir présidé. Ce fut un réel honneur pour moi de tous vous avoir dans mon jury.

Je tiens à saluer ici l'équipe *cosmologie observationnelle* du LAM dans laquelle je me suis très vite senti à l'aise. Je citerai en particulier, et par ordre alphabétique Christophe

Adami (merci encore pour la relecture de ce manuscrit), Stéphane Arnouts, Vincent Le Brun, Christian Marinoni, Alain Mazure, Laurence Tresse.

Tous font partie de la *VVDS team* que je n'oublie pas et qui a permis à ce projet de voir le jour, lancé voici maintenant 10 ans en 1995 par une poignée d'entre eux. L'équipe compte dorénavant une quarantaine de chercheurs. Je rajouterai à ceux déjà cités des personnes comme Bianca Garilli, Dario Maccagni, Fabrice Lamareille, Henry McCracken, Sébastien Foucaud (mon prédécesseur, aussi bien en thèse avec Olivier, qu'en prépa MP avec Coup's à Camille-Guérin !), Stéphane Paltani et Thierry Contini. Et bien sur, Agnieszka Pollo, avec qui j'ai énormément apprécié de travailler sur le calcul de la fonction de corrélation. J'en profite pour saluer Jérémy Blaizot sur qui j'ai mis beaucoup de pression pour qu'il me construise ces 50 simulations et les mette en ligne sur la base donnée *GallCS*.

Je remercie bien sur les petites mains du LAM qui font tourner la boutique en s'occupant de l'informatique (Chrystel, Thomas), de la logistique (Alain, Babeth, Christophe, Domi, Hélène, Jean-Antoine, Nathalie, Odile, Phiphi, Raphaël), de la photographie (Jean-Pierre), sans oublier la cuisine (Corinne) !

Je remercie aussi chaleureusement Valérie Kazarian et Valérie Zinck de l'IUT, Jean-Marc Pislott, Vincent Le Brun et Remi Morin à l'Université de Provence qui m'ont permis de passer de l'autre côté du bureau et ainsi d'encadrer quelques travaux pratiques et travaux dirigés en informatique, thermodynamique et chimie.

Je salue tous les étudiants, stagiaires, thésards, et CDDs qui ont rendu l'ambiance au labo très agréable : les *boulets* Sébastien V, Sébastien L, Olivier (alias Skyp), Sébastien H (Mister Chocolat), Claire (une partie de tarot ?), Delphine et Marie-Hélène (déstress !!), la *Lascot Team* avec Fabien (et sa petite famille avec Agnès et Mathieu), Guillaume, Michael et Yannick, et enfin Alexie (mon ex-voisine à l'internat en Math Spé 3/2, le monde est petit !!), Céline, Fabrice, Lidia, Mélanie, Sébastien R, Sylvain (reprends bien le flambeau) et Thomas.

Je n'oublie pas les nombreux étudiants que j'ai eu l'occasion de rencontrer lors de mes deux séjours de 3 semaines à Merate. Je ne les citerai pas tous. Je pense par exemple à Flora, Igor, Luca, Lucia, Paolo, Paolino, Stephania, Vincenzo, Alberto, Giorgio, Filiberto, Nino, ... Nous avons partagé plusieurs repas en commun à la *Foresteria* et il n'y a pas à dire, mais les pâtes que l'on mange en Italie sont vraiment les meilleurs !!

La vie hors du labo a elle aussi été très agréable. Le mérite en revient en grande partie à mes collègues de (ex-)Sup'Phy(-future Centrale Marseille à ce qu'on dit) qui sont restés à Marseille, eux aussi pour faire une thèse et avec qui je n'ai eu que de bons moments. Les deux semi-marathoniens (Marseille-Cassis) David et Bertrand (je viendrais te voir à Tours, c'est promis, cher ex-binôme de TP, et avec un pichet de rosé s'il te plaît !!), Hélène (qui nous a fait découvrir le potimarron. Excellent !!), Laetitia & Lewis (les ex-nouveau grenoblois qui ont inventé le *Glouphytendeur*, à découvrir sur leur blog *LaëLew*¹), Yann (sans

¹voir <http://laelew.canalblog.com>

qui le club-astro de Sup'Phy n'aurait jamais été ce que nous en avons fait), et Cédric (le savoyard qui ne manque jamais ni l'ouverture, ni la fermeture d'une saison de ski, et qui participe aux sorties astro toujours habillé uniquement d'un t-shirt, même en plein hiver !!)

Je remercie toute ma très grande famille (suffisamment grande pour faire un VVDS-bis !!) et en particulier mes parents, Catherine et Jacques, mon frère Sylvain, ma marraine Catherine et ma grand-mère Alice. Merci à tous pour votre soutien et l'intérêt que vous avez accordé à mes études au sens large. Je sais que vous n'avez pas fini de me poser des questions (surtout si vous lisez tout ou partie des pages qui suivent . . .), mais je continuerai à essayer au mieux d'y répondre. Promis !

Enfin, je ne peux pas terminer ces remerciements sans dire un mot sur ma tendre et adorable Anne au magnifique sourire. Sans toi, rien n'aurait été pareil . . .

*La Terre est le berceau de l'Humanité.
Mais on ne passe pas sa vie dans un berceau.*

Konstantin E. Tsiolkovsky (1857-1935)

Table des matières

Remerciements	i
Table des matières	v
Introduction	1
1 Distribution des galaxies à grande échelle	3
1.1 Scénario et modèles cosmologiques	3
1.1.1 Cosmogonie	3
1.1.2 Les modèles d'Univers	4
1.2 Formation des grandes structures de l'Univers	6
1.2.1 Modèles hiérarchiques	6
1.2.2 Concept et notion de biais	7
1.3 Observation des grandes structures	7
1.3.1 Une vision multi-longueur d'onde des galaxies	7
1.3.2 Une vision tridimensionnelle de l'Univers	9
1.3.3 Structures dans l'Univers proche ($z \lesssim 0.3$)	10
1.3.4 Structures dans l'Univers lointain ($z \gtrsim 0.3$)	14
1.4 La fonction de corrélation à 2 points : un outil majeur	15
1.4.1 Un outil pour quantifier la distribution des galaxies et de la matière : définition et estimateur	15
1.4.2 Fonction de sélection et comparaison des mesures	17
1.4.2.1 Population globale	18
1.4.2.2 Sélection en fonction de la luminosité absolue	18
1.4.2.3 Sélection en fonction du type	18
1.4.2.4 Galaxies à discontinuité Lyman (LBG)	22
1.4.2.5 Galaxies EROs	23
1.4.3 Fonction de corrélation et biais	24
1.5 État des lieux : mesures de l'agrégation des galaxies	25
1.5.1 Dans l'Univers proche	25
1.5.1.1 Population globale	25
1.5.1.2 Les amas	27
1.5.1.3 En fonction de la luminosité	28
1.5.1.4 En fonction du type morphologique	31
1.5.1.5 En fonction de la couleur/type spectral	32
1.5.1.6 Synthèse	38
1.5.2 À grand décalage spectral	39

1.5.2.1	Population globale	39
1.5.2.2	Les galaxies à discontinuité Lyman	40
1.5.2.3	Les galaxies extrêmement rouges	42
1.6	Évolution des grandes structures : prévision des modèles	43
1.7	Synthèse	47
2	Le sondage profond VIMOS-VLT <i>Deep Survey</i>	49
2.1	Des sondages d'Univers lointain	49
2.2	Le sondage spectroscopique profond VVDS	51
2.2.1	But et stratégie	51
2.2.2	L'instrument VIMOS	55
2.2.2.1	Imagerie et spectroscopie multi-objet	55
2.2.2.2	Spectroscopie intégrale de champ	56
2.2.3	Observations	56
2.3	Réduction des données spectroscopiques et détermination des <i>redshifts</i>	58
2.3.1	Réduction des données	58
2.3.2	Détermination des <i>redshifts</i>	61
2.3.2.1	Raies en émission, en absorption	61
2.3.2.2	Méthodologie	65
2.3.2.3	Indices de confiance	66
2.3.2.4	Exemple de spectres	70
2.4	Première campagne d'observation : les données	73
2.4.1	Les champs <i>Deep</i>	73
2.4.1.1	Le VVDS-02h	73
2.4.1.2	Le <i>Chandra Deep Field South</i>	75
2.4.2	Les champs <i>Wide</i>	77
3	Agrégation des galaxies : outil de mesure et erreurs associées	81
3.1	La fonction de corrélation spatiale	81
3.1.1	$\xi(r_p, \pi)$ à 2 dimensions	81
3.1.2	La projection $w_p(r_p)$	84
3.2	Contraintes observationnelles propres au VVDS	85
3.2.1	Les masques photométriques	85
3.2.2	Le critère de sélection observationnelle	85
3.2.3	La sélection des cibles	85
3.2.4	L'incertitude sur la mesure du <i>redshift</i>	85
3.2.5	L'incomplétude spectroscopique	87
3.3	$\xi(r_p, \pi)$: l'outil de mesure	87
3.3.1	Le catalogue de points générés aléatoirement	88
3.3.2	L'attribution d'un poids à chaque galaxie	89
3.3.3	Le calcul des distances et l'échantillonnage	92
3.4	Mesure des paramètres r_0 et γ	93
4	Calcul des erreurs : utilisation des simulations <i>GalICS</i>	97
4.1	Simulations numériques	98
4.1.1	Hubble-Volume	98
4.1.2	<i>GalICS</i>	100

4.2	Création de catalogues fictifs	101
4.2.1	Les étoiles	103
4.2.2	Les rayons apparents de chaque objet	103
4.2.3	Comparaison des catalogues - propriétés	106
4.3	Comparaison des résultats obtenus en appliquant ou non la stratégie d'observation du VVDS	108
4.4	Estimation des erreurs de mesures	110
4.5	<i>Article I : The VIMOS VLT Deep Survey - Computing the two point correlation statistics and associated uncertainties</i>	<i>112</i>
5	Agrégation des galaxies dans le sondage VVDS	127
5.1	Évolution de la fonction de corrélation globale $\xi(r)$ à $l_{AB} = 24$ jusqu'à $z \sim 2$	127
5.1.1	Données utilisées	127
5.1.2	Résultats	129
5.1.3	<i>Article II : The VIMOS VLT Deep Survey - The evolution of galaxy clustering to $z \sim 2$ from first epoch observations</i>	<i>130</i>
5.2	Évolution de la fonction de corrélation par type de galaxies jusqu'à $z \sim 2$. .	140
5.2.1	Définition des types	140
5.2.2	Mesures et erreurs associées	141
5.2.3	Résultats	145
5.2.4	<i>Article III : The VIMOS VLT Deep Survey - the evolution of galaxy clustering per spectral type to $z \sim 1.5$</i>	<i>149</i>
5.3	Évolution de la fonction de corrélation pour un échantillon de galaxies sélectionnées dans l'ultraviolet	162
5.3.1	Galex : observation en NUV et FUV	162
5.3.2	Fonction de corrélation en NUV	165
5.3.3	Fonction de corrélation en FUV	166
5.4	Évolution du biais non-linéaire jusqu'au <i>redshift</i> $z=1.5$	168
5.4.1	Résultats	168
5.4.2	<i>Article IV : The VIMOS VLT Deep Survey - evolution of the non-linear galaxy bias up to $z=1.5$</i>	<i>169</i>
6	Évolution de l'agrégation des galaxies depuis $z \sim 2$: une nouvelle référence	199
6.1	Référence observationnelle	199
6.2	Implication sur l'évolution des galaxies	208
	Conclusions et perspectives	209
	A Correlation Function Tool - release xi_v0.9.8	211
	Bibliographie	225
	Liste des tables	231
	Liste des figures	233

Introduction

L'homme a toujours été fasciné par l'origine de l'Univers, sa formation, jusqu'à l'apparition de la vie sur terre. Désirant se positionner dans l'immensité et chercher sa place parmi toutes les beautés du ciel nocturne, de la même manière que sa curiosité l'a poussé à parcourir le monde à la recherche de nouveaux continents, l'homme s'est mis à cartographier, répertorier et analyser les objets qui remplissent sa vision. Si les civilisations successives ont attribué des noms et des sens différents à chaque partie du ciel, il n'en reste pas moins que parallèlement progressaient la description et la connaissance de l'Univers observable.

Au 20^{ième} siècle naît la cosmologie, science qui consiste à étudier l'Univers et son évolution dans son ensemble. Comme toute science, ses découvertes, observationnelles ou théoriques la font évoluer jusqu'à lui rendre la possibilité de faire des prédictions.

Depuis plus d'une décennie, les avancées technologiques ont permis d'augmenter la surface collectrice des télescopes, la résolution spatiale et la sensibilité des détecteurs. Le développement des capacités de traitement, d'analyse, d'archivage et de circulation de l'information sont allés de pair permettant l'observation du ciel avec toujours plus de détails. La cosmologie observationnelle a ainsi réalisé de grandes avancées dans la compréhension des différentes composantes astrophysiques, depuis le proche environnement de notre Système Solaire, de notre Galaxie, jusqu'aux fluctuations du fond cosmologique, en passant par l'émission des galaxies et des amas.

Un des domaines en pleine mutation concerne l'étude de la formation des grandes structures que l'on observe dans l'Univers. Dans un Univers en expansion, les structures se forment (galaxies, amas de galaxies et super amas) et évoluent selon leur masse et en fonction de leur environnement. L'observation et l'étude de ces structures sont importantes pour comprendre où et comment se forment les galaxies. Il est aussi nécessaire de pouvoir "remonter" la flèche du temps pour ainsi retracer l'évolution de ces structures et pouvoir dater les époques importantes de formation des galaxies.

Pour cela, de grands sondages du ciel sont réalisés. Leur but est d'observer un très grand nombre de galaxies, et de mesurer leurs distances, ou encore de les positionner dans le temps, pour reconstituer ces structures à trois dimensions. Le sondage *VIMOS-VLT Deep Survey* (VVDS) se propose justement d'ajouter sa pierre à cet édifice pharaonique en observant les premières structures de l'Univers, telles qu'elles étaient quand celui-ci n'était âgé que d'environ un milliard d'années. L'analyse des galaxies observées permet de sonder l'Univers sur près de 90% de son âge et ainsi de suivre la formation et l'évolution des structures. C'est la première fois qu'une étude sur une si grande plage de temps est réalisée, et avec un minimum d'a priori sur les galaxies observées, les résultats n'en étant alors que plus représentatif des propriétés de l'Univers.

Ainsi, dans le cadre de cette thèse, au sein du consortium VVDS, j'ai été amené à mesurer le *redshift* de plusieurs milliers de galaxies. J'ai développé des outils d'analyse pour

quantifier l'évolution des structures formées par les galaxies à différentes époques, pour chiffrer comment ces galaxies, en fonction de leurs propriétés physiques, sont distribuées et réparties dans l'Univers. Dans un premier temps, je décrirai les connaissances acquises ces dernières années par la communauté scientifique internationale sur les propriétés d'agrégation des galaxies, essentiellement dans l'Univers local, mais aussi les quelques mesures disponibles à grands *redshifts*. Je décrirai ensuite le sondage VVDS, puis les outils et méthodes d'analyse des données que j'ai développé. Je présenterai enfin les résultats obtenus au cours de cette thèse sur l'évolution de l'agrégation des galaxies qui font office maintenant de nouvelles références en la matière.

Chapitre 1

Distribution des galaxies à grande échelle

Résumé

1.1 Scénario et modèles cosmologiques	3
1.2 Formation des grandes structures de l'Univers	6
1.3 Observation des grandes structures	7
1.4 La fonction de corrélation à 2 points : un outil majeur	15
1.5 État des lieux : mesures de l'agrégation des galaxies	25
1.6 Évolution des grandes structures : prévision des modèles	43
1.7 Synthèse	47

Ce chapitre présente le contexte scientifique dans lequel s'est déroulée cette thèse. Je présente tout d'abord les modèles cosmologiques et mécanismes de formation des grandes structures observées dans l'Univers. J'introduis ensuite les observations de ces structures et la connaissance qualitative que nous en avons dans l'Univers local ainsi que dans l'Univers lointain. Je présente ensuite l'outil statistique, la fonction de corrélation spatiale, qui nous permet de quantifier ces mêmes structures, et les mesures existantes dans l'Univers proche et lointain. Je termine par un tour d'horizon des prédictions théoriques disponibles aujourd'hui sur l'agrégation des galaxies.

1.1 Scénario et modèles cosmologiques

1.1.1 Cosmogonie

Voici 2300 ans, Epicurus (341-270 avant J.-C.) se demandait si d'autres mondes identiques et/ou différents du notre pouvaient exister quelque part ailleurs, cachés derrière les étoiles, alors qu'Aristote (384-322 avant J.-C.) enseignait qu'il n'en existait qu'un. De longs débats ont eu lieu depuis quant à la nature de l'Univers et notre position dans celui-ci, les arguments s'appuyant sur les connaissances de l'époque, mais jamais traduites en terme de physique.

Voici plus de 200 ans, une seule galaxie était considérée. L'idée de galaxie venait d'être lancée par Thomas Wright en 1750. En 1785, William Herschel en dessina une carte en considérant le soleil comme proche de son centre. Ainsi quelques années à peine après la réhabilitation de Galilée (1757) et du système héliocentrique de Copernic dans lequel toutes les planètes tournent autour du soleil, un nouveau débat s'installe : celui de la place du soleil. Alors que Kant croit en l'existence d'autres "Univers-îles" à l'extérieur de la Voie Lactée, l'idée d'une seule galaxie persiste jusqu'au début du 20^{ième} siècle avec l'Univers de Kaypen. Cette controverse culmine avec le grand débat entre Curtis et Shapley, partisan d'une seule super galaxie. Elle fut close par Edwin P. Hubble en 1928 qui mesura la distance de nébuleuses grâce aux étoiles variables Cépéïdes dont la relation luminosité-distance fut établi par Henrietta Leavitt peu auparavant en 1912.

À partir de là, le décalage vers les grandes longueurs d'onde (*redshift*) du spectre émis par une galaxie qui s'éloigne étant connu, Hubble établit la loi qui porte son nom, à savoir que la vitesse de récession des galaxies est proportionnelle à leur distance (d'un facteur H mesuré pour la première fois par Hubble égal à $540 \text{ km.s}^{-1}.\text{Mpc}^{-1}$). Toutes les galaxies semblent s'éloigner. Hubble observe un Univers en expansion. Il note aussi que cet Univers semble homogène et isotrope. Mais les premiers sondages systématiques conduisent à la découverte de nombreux amas, et même groupes d'amas de galaxies. Les premières idées de modélisation de l'Univers pouvaient être lancées.

1.1.2 Les modèles d'Univers

Parralèlement à ces observations, un cadre théorique se met en place. Tout d'abord, il y a un siècle, Albert Einstein découvre la relativité. En 1915, il publie la théorie géométrique de la relativité générale. Il développe un modèle d'Univers et introduit le *principe cosmologique*. Celui-ci dit que l'Univers est homogène et isotrope et le temps universel. Ceci permet de définir une métrique dans l'espace-temps et d'établir des relations théoriques, les équations d'Einstein, entre la courbure de l'Univers et son contenu en énergie-matière.

La métrique la plus générale est celle de Friedmann-Lemaître-Robertson-Walker et prends la forme suivante (Peacock, 1999) :

$$ds^2 = c^2 dt^2 - R^2(t) [f^2(r) dr^2 + g^2(r) d\phi^2] \quad (1.1)$$

Les distances sont décomposées en un produit entre un facteur d'échelle $R(t)$ dépendant du temps et une coordonnée comobile r indépendante de celui-ci.

Dans le cas de la loi de Hubble introduite précédemment, le paramètre du même nom qui quantifie l'expansion s'exprime en fonction du facteur d'échelle par

$$H(t) \equiv \frac{\dot{R}(t)}{R(t)} \quad (1.2)$$

La constante de Hubble H_0 correspond à la valeur actuelle de la vitesse d'expansion de l'Univers. Afin de s'affranchir de l'incertitude sur cette valeur, on introduit la constante sans dimension h définie par

$$h = \frac{H_0}{100 \text{ km.s}^{-1}.\text{Mpc}^{-1}} \quad (1.3)$$

Les modèles d'Univers qui se développent ainsi au cours du début du 20^{ième} siècle font intervenir d'autres grandeurs cosmologiques.

- le facteur de décélération q_0 , qui, selon sa valeur, décrira un Univers en expansion ($q_0 > 1/2$), en contraction sur lui-même ($q_0 < 1/2$), ou bien stationnaire ($q_0 = 1/2$). Il s'exprime par

$$q_0 = - \left(\frac{\ddot{R}}{\dot{R}^2} \right)_{t_0} = - \frac{\ddot{R}(t_0)}{H_0^2} \quad (1.4)$$

- le paramètre de densité, Ω_m , défini comme le rapport entre la densité actuelle d'énergie-masse de l'Univers ρ_0 et la densité critique $\rho_c = 3H_0^2/8\pi G$, où G est la constante universelle de la gravitation
- la constante cosmologique Λ qui intervient dans l'expression de la densité d'énergie du vide $\Omega_\Lambda = \Lambda/3H_0^2$
- un facteur de courbure d'espace κ

Dans le modèle d'Einstein-de Sitter, on prend par exemple $\Omega_m = 1$ et la constante cosmologique nulle. Le modèle privilégié en ce début de 21^{ème} siècle, dit modèle de “concordance” est un Univers plat ($\Omega_m + \Omega_\Lambda = 1$) avec une constante cosmologique non nulle. Les données obtenues à partir de la sonde WMAP (*Wilkinson Microwave Anisotropy Probe*) concluent à un Univers plat dominé par une constante cosmologique avec $\Omega_m = 0.27 \pm 0.04$ et $h = 0.72 \pm 0.05$ pour un âge de 13.7 ± 0.2 milliards d'années (Spergel *et al.*, 2003). Ces mesures sont en accords avec des analyses effectuées sur des supernovae ou encore les amas de galaxies (Bahcall *et al.*, 2003). La figure 1.1 illustre, dans le cas de trois cosmologies différentes où $H_0 = 70 \text{ km.s}^{-1}.\text{Mpc}^{-1}$, le temps pris par la lumière émise par une galaxie située à un *redshift* z pour nous parvenir.

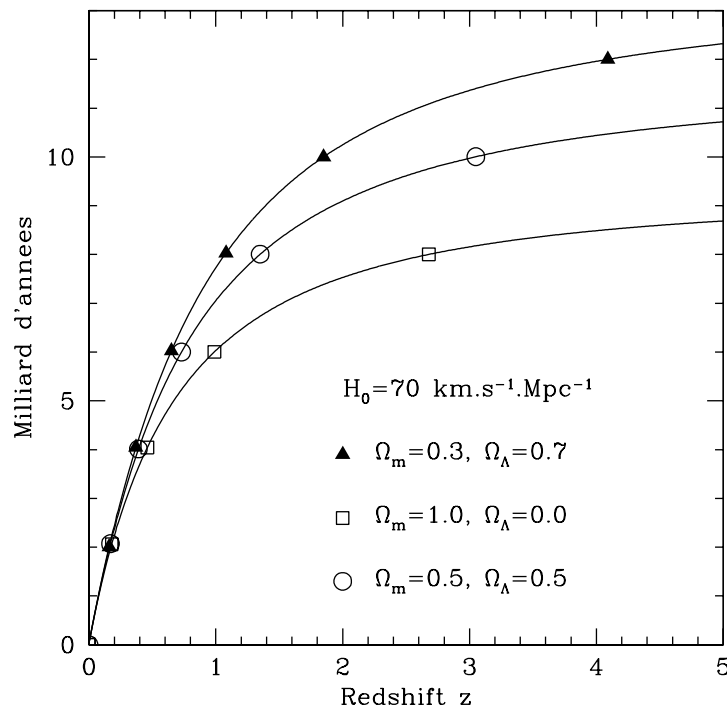


FIG. 1.1 – Relation temps-*redshift*

1.2 Formation des grandes structures de l'Univers

1.2.1 Modèles hiérarchiques

Plusieurs facteurs importants supportent notre vision actuelle de la formation des structures dans l'Univers. En voici quelques exemples :

- la découverte par Hubble en 1928 d'une relation linéaire entre la vitesse et la distance des galaxies, relation interprétée comme le résultat de l'expansion de l'Univers au travers des équations de Friedmann-Lemaître qui décrivent un Univers homogène et isotrope.
- la découverte par hasard en 1965 par Penzias et Wilson du rayonnement cosmologique (*Cosmic Microwave Background* - CMB) à 2.728 ± 0.004 K (Spergel et al., 2003), décrit d'un point de vue théorique dès 1940 par Gamow. Le CMB correspond au rayonnement émis suite au découplage avec la matière lorsque les électrons et protons se sont recombinés de part le refroidissement de l'Univers du à son expansion.
- l'observation en 1992 par le satellite COBE (*Cosmic Background Explorer*) de ce dernier. La découverte de fluctuations de température, de l'ordre de $\Delta T/T \sim 10^{-5}$, sur tout le ciel, apportant des informations importantes sur les conditions initiales de formation des structures.

Avant d'aller plus loin, il est important de rappeler que l'étude directe de la distribution de la matière dans l'Univers est celle de la matière visible. Près de 83% de la matière réellement présente nous échappe ($\Omega_b/\Omega_m \simeq 0.17$, Spergel et al. (2003)). Heureusement, il est possible d'étudier cette matière noire par l'influence gravitationnelle qu'elle joue sur les structures, l'expansion de l'Univers, et courbes de vitesse des galaxies. Elle a en effet été mise en évidence par l'étude des courbes de rotation des galaxies pour lesquelles la vitesse reste constante et importante en périphérie de la structure visible de la galaxie alors qu'elle devrait diminuer si l'on ne tient compte que de la masse visible (lois de Kepler). À l'échelle des amas, la dispersion de vitesse des galaxies et la distorsion des galaxies d'arrière plan par effet de lentille gravitationnelle trahit la présence d'une matière massive invisible.

On retiendra deux sortes de matière noire : l'une froide qui intervient dans les modèles CDM (pour *Cold Dark Matter*) et l'autre chaude dans les modèles HDM (pour *Hot Dark Matter*). La "température" de cette matière invisible donne une image de la vitesse et de la masse des particules qui la constitue.

Dans un modèle CDM, les fluctuations initiales du champ de densité (dont le CMB est une représentation) se sont condensées en petites structures sous l'effet de la gravitation. Et pour les mêmes raisons, elles se sont effondrées à leur tour pour former des halos de plus grandes masses, lieux de formation des galaxies. C'est le modèle "hiérarchique". Dans un modèle HDM, ce sont les grandes structures initiales qui se sont subdivisées. C'est le modèle "monolithique".

Actuellement, le scénario privilégié est celui du CDM muni d'une constante cosmologique. Dans ce modèle, les fluctuations primordiales sont supposées gaussiennes et adiabatiques et son spectre de puissance $P(k)$ peut être calculé. Les gros halos de matière se forment par fusion de halos plus petits. Les plus grosses structures non-linéaires actuelles sont les amas de galaxies. Les filaments et super-amas sont eux encore à un

stade de croissance linéaire. L'observation de galaxies à grand *redshifts* semble favoriser un tel scénario (voir partie 1.5).

1.2.2 Concept et notion de biais

Jusqu'au début des années 1980, l'hypothèse généralement faite était que les galaxies tracent le champ de distribution de la masse présente dans l'Univers. Puisque ce champ est continu, les galaxies représentent alors un échantillonnage poissonnien de celui-ci. Dans un volume V , le nombre de galaxies observées est alors proportionnel à la masse présente.

Or on observe que le regroupement des amas, mesuré au moyen de la fonction de corrélation spatiale (voir chapitre 1.5.1.2) est plus important que celui des galaxies. Les amas sont ~ 12 fois plus corrélés que les galaxies. La distribution à grande échelle de la matière n'est donc pas tracée de la même façon par les amas et les galaxies. Et rien ne dit que des amas ou des galaxies suivent le mieux le champ de densité.

Ce fut la raison principale qui amena Kaiser (1984) à introduire le concept de biais et la théorie des pics. La formation des galaxies et des amas ne pouvant être le résultat de conditions initiales particulières différentes, les amas ont dû se former là où, localement, les fluctuations du champ de densité étaient très importantes, au-delà d'un seuil critique. Cela explique ainsi une amplitude de corrélation des amas plus élevée que celle de la matière noire. D'autres arguments en faveur d'une distribution biaisée des galaxies sont par exemple la dépendance de l'amplitude de la fonction de corrélation des galaxies en fonction de leur nature (luminosité (section 1.5.1.3), type morphologique (section 1.5.1.4), couleur, type spectral (section 1.5.1.5)).

Une description plus formelle de la notion de biais sera introduite et développée dans la section 1.4.3.

1.3 Observation des grandes structures

Une science comme la cosmologie observationnelle, comme presque toutes les autres d'ailleurs, dépend des avancées technologiques de son époque. Ainsi, chaque nouvelle génération de télescope apporte son lot de nouvelles découvertes. Depuis les plaques photographiques jusqu'aux capteurs CCD, des télescopes de ~ 1 mètre de diamètre aux géants Keck et VLT de ~ 10 mètres, l'observation du ciel s'est faite de plus en plus profonde, dans presque toutes les longueurs d'onde, répertoriant ainsi le contenu de l'Univers.

1.3.1 Une vision multi-longueur d'onde des galaxies

Le premier outil utilisé pour enregistrer des images du ciel fut la plaque photographique. Elle fut d'une grande utilité aussi bien pour la recherche de comètes dans le système solaire que pour la cartographie des étoiles, et le recensement de nébuleuses ou galaxies. À chaque objet ainsi répertorié est associée une position (astrométrie) et une magnitude (photométrie) qui permet de classer les objets des plus brillants aux plus faibles. Or quelque soit l'objet étudié, son flux lumineux reçu dépend de la longueur d'onde à laquelle

on l'observe. Ceci est une conséquence directe de ses propriétés physiques (température, composition chimique, vitesse relative de déplacement, ...).

Ainsi, pour ce qui concerne l'étude des galaxies, les observations modernes du ciel se font maintenant à différentes longueurs d'onde, au travers de filtres de bandes passantes larges ou étroites. Ces sondages multi-couleur ont réellement débuté au début des années 1980. L'analyse du flux infrarouge émis par les galaxies permet de quantifier la masse d'étoiles évoluées, alors que celui émis à plus basses longueurs d'onde, dans le bleu et ultraviolet permet d'étudier la formation stellaire.

Grâce aux avancées de la technique, les plaques photographiques ont laissé la place aux caméras numériques CCDs (*Charged-Coupled Device*). L'utilisation combinée de ces nouveaux détecteurs avec des télescopes de diamètre de plus en plus grand (première lumière du télescope Canada - France - Hawaï de 3.6 mètres de diamètre en 1979, télescopes de l'ESO à La Silla accessibles aux astronomes français, et toute une batterie de télescopes au niveau mondial) permet d'observer le ciel très profondément. Les magnitudes alors atteintes furent de $B \sim 25$. Là où les premiers sondages se focalisaient sur l'étude de l'Univers local, l'observation de l'Univers lointain dans différentes bandes spectrales fut rendu possible par ces nouvelles techniques.

L'utilisation de caméras à grand champ aida considérablement à la réalisation de grands sondages du ciel. Le champ de vue de la caméra CFH8k était par exemple de $29' \times 29'$, soit l'équivalent de la pleine lune. Elle fut utilisée dans le cadre du *Canada France Deep Field* où ~ 100000 galaxies furent observées dans un champ de un degré carré jusqu'à une magnitude $I_{AB} = 25.5$ (McCracken *et al.*, 2001). La caméra CFH12k qui couvre un champ de $42' \times 28'$ fut utilisée pour réaliser la partie photométrique du sondage *VIMOS-VLT Deep Survey* (voir chapitre 2). Actuellement, la caméra MEGACAM dont le champ de vue fait 1 degré carré est utilisée dans le cadre du relevé photométrique *CFHT Legacy Survey*¹. La partie imagerie du sondage SDSS est réalisée avec une caméra dont le champ fait 2.5 degrés carrés. Celle du 2dFGRS est faite à partir de plaques photographiques digitalisées du sondage Stromlo-APM. Le télescope de 2.2m de l'ESO à La Silla, utilisé pour des observations en U dans le VVDS ainsi que dans le cadre du sondage COMBO-17, est muni de la caméra WFI (*Wide Field Imager*) dont le champ fait $33' \times 34'$. Enfin, on pourra citer aussi le télescope Isaac Newton aux Canaries muni d'une caméra à grand champ ($\sim 0.3 \text{ deg}^2$) ainsi que le télescope UKIRT à Hawaï qui peut réaliser dans l'infrarouge des images dont le champ fait $\sim 0.9 \text{ deg}^2$. Ce dernier est utilisé dans le cadre du sondage UKIDSS. Les contraintes des observations faites depuis le sol sont essentiellement liées aux conditions météorologiques et à la qualité du *seeing*. Ce dernier correspond à la résolution angulaire maximale des images obtenues. La résolution est en effet dégradée par les turbulences de l'atmosphère.

Afin d'améliorer la résolution des images produites depuis le sol, le télescope spatial *Hubble* (HST) fut lancé en 1990. La résolution depuis l'espace ne dépend que de la limite de diffraction de l'instrument qui est proportionnelle à λ/D , avec D le diamètre du miroir du télescope et λ la longueur d'onde à laquelle on observe. Le HST a permis la réalisation des très profonds sondages ($I_{AB} \sim 27.6$) *Hubble Deep Field North* et *South* (HDF-N et HDF-S) mais qui couvrent des champs de vue de seulement ~ 4 minutes d'arc carrées. Même si le nombre d'objets observés reste ainsi limité (~ 2500), la profondeur fait que des objets à très grands *redshifts* ($z \sim 6$) se retrouvent dans ces échantillons. La résolution

¹voir <http://www.cfht.hawaii.edu/Science/CFHTLS>

des images du HST permet d'étudier les propriétés morphologiques des galaxies jusqu'à des décalages spectraux au-delà de $z=1$. Ainsi, pour l'un des derniers sondages majeurs réalisés par le HST, 640 orbites ont été allouées pour l'observation du champ COSMOS qui couvre 2 degrés carrés avec la caméra *Advanced Camera for Survey*.

Les relevés photométriques du ciel, même s'ils doivent trouver un compromis entre surface du ciel observée et profondeur atteinte, sont généralement tous multi-couleur. Il est possible aujourd'hui d'observer une zone du ciel à de très nombreuses longueurs d'onde avec par exemple les instruments suivants :

- rayons X avec les satellites XMM-Newton et Chandra
- ultraviolet proche et lointain avec le satellite GALEX
- ultraviolet, visible et proche infrarouge avec les télescopes au sol
- infrarouge lointain avec le satellite Spitzer

De manière générale, la réalisation d'un sondage photométrique doit faire l'objet d'un compromis entre la surface du ciel couverte et la profondeur que l'on souhaite atteindre.

1.3.2 Une vision tridimensionnelle de l'Univers

La photométrie n'est pas capable de donner directement une information précise sur la distance des sources observées. La spectroscopie est alors utilisée pour mesurer précisément le décalage vers le rouge et positionner les objets observés dans l'espace local à trois dimensions. En observant la distribution 3D à différentes époques on a ainsi une vision de l'évolution des structures. La troisième dimension (*redshift* z) est définie comme $z = \lambda/\lambda_0 - 1$ où λ_0 est la longueur d'onde à laquelle une onde lumineuse est émise par une galaxie et λ celle à laquelle elle est reçue. La détermination de z nécessite donc d'enregistrer le spectre reçue d'une galaxie.

Le but principal des sondages de galaxies avec mesure du *redshift* est de cartographier l'Univers à la fois dans l'espace physique (coordonnées angulaires) et dans l'espace des vitesses (en particulier l'écart avec l'expansion uniforme de Hubble) afin de comprendre l'agrégation et la dynamique des galaxies. De là il est possible de déduire des informations sur la distribution de la matière, celle de la lumière et leur relation. On est alors capable de quantifier le biais entre la distribution de masse totale et celle visible. De plus, l'information spectroscopique fournit des renseignements sur le contenu stellaire et gazeux des galaxies.

Une avancée technologique majeure a permis d'envisager et de réaliser maintenant de très grands sondages. La technique multi-objet consiste à enregistrer simultanément le spectre de plusieurs galaxies. Les instruments multi-fibre sont utilisés pour la mesure des spectres de galaxies proches, alors que les instruments multi-fente permettant une correction plus précise du fond de ciel sont utilisés pour les galaxies distantes. Deux grands sondages de galaxies locales, le 2dFGRS et SDSS (partie 1.3.3), de part leur ampleur, ont permis des avancées considérables dans la description de l'Univers local. Un des premiers spectrographes multi-fente pour l'observation des galaxies distantes, MOS-SIS, fut installé au télescope CFHT (Le Fèvre *et al.*, 1994a) et utilisé dans le cadre du sondage spectroscopique *Canada France Redshift Survey* (Lilly *et al.*, 1995b; Le Fèvre *et al.*, 1995, CFRS). Le CFRS (partie 1.3.4) a établi pour la première fois une référence sur l'évolution des galaxies en remontant jusqu'à la moitié de l'âge de l'Univers ($z \sim 1.2$). Des sondages comme le 2dFGRS ou SDSS couvrent de très grands champs, sur presque la moitié du

ciel. Mais la réalisation de sondages spectroscopiques de l'Univers lointain nécessite des images très profondes, et donc des champs beaucoup plus petits. Les objets étant faibles les temps de poses (en imagerie et spectroscopie) augmentent en conséquence.

Le CFRS a eu un rôle précurseur en démontrant la faisabilité technique de sondages spectroscopiques profonds. Ainsi d'autres projets lancés ces dernières années se sont concentrés sur l'étude de l'Univers à des *redshifts* intermédiaires de l'ordre de $z \sim 0.5 - 1$ comme par exemple le *Canadian Network for Observational Cosmology* (Yee *et al.*, 2000, CNOC2). Deux exemples de grands sondages en cours pour étudier l'Univers à grand *redshift* sont le *Deep Extragalactic Evolutionary Probe* (Davis *et al.*, 2003, DEEP2) et le *VIMOS-VLT Deep Survey* (Le Fèvre *et al.*, 2005c, VVDS, partie 2).

Malgré la technique multi-objet, un compromis doit être trouvé entre temps de pose, profondeur et taille du champ observé. La mesure des *redshifts* est très précise mais l'échantillonnage n'atteint pas 100%. L'information multi-longueur d'onde permet de déterminer le *redshift* photométrique de presque tous les objets d'un catalogue. Cette méthode qui repose sur le flux émis par les galaxies à différentes longueurs d'onde à l'avantage de pouvoir attribuer une distance à plusieurs milliers (millions) d'objets. L'étude des propriétés des galaxies peut se faire alors avec une excellente statistique. Mais l'inconvénient de cette méthode est qu'elle est peu robuste et beaucoup moins précise que celle des *redshifts* spectroscopiques. Dans le cadre du HDF-N et HDF-S, 7 filtres ont été utilisés. Ils couvrent une plage en longueur d'onde de 3000Å à 22200Å permettant d'atteindre des *redshifts* de l'ordre de $z \sim 5$ (Arnouts *et al.*, 1999). Le sondage *Classifying Objects by Medium-Band Observations* (COMBO-17) a consisté à observer un champ de 0.78 degrés carrés dans 17 filtres étroits afin de bien échantillonner le spectre, à savoir le flux émis en fonction de la longueur d'onde, par chaque galaxies. La précision atteinte est de $dz/(1+z) \sim 0.02$ (Wolf *et al.*, 2004) pour un domaine de validité en *redshift* de $[0.2 - 1.2]$, limité par l'absence d'observation en infrarouge. D'autres sondages utilisent cette technique comme par exemple le *Calar Alto Deep Imaging Survey* (Phleps et Meisenheimer, 2003, CADIS) et le *Great Observatories Redshift Survey* (Giavalisco *et al.*, 2004, GOODS). On retiendra que l'erreur associée à un *redshift* photométrique reste 100 fois supérieure à celle d'un *redshift* spectroscopique. D'où le développement ces dernières années des sondages spectroscopiques.

1.3.3 Structures dans l'Univers proche ($z \lesssim 0.3$)

Alors que les observations montraient déjà des régions de l'Univers avec des amas et des surdensités de galaxies observés par Huchra *et al.* (1983) dans le sondage *Center for Astrophysics survey* (CfA) et Giovanelli et Haynes (1985) dans le *Perseus-Pisces survey*, la première avancée majeure dans l'identification de la distribution complexe des galaxies en grandes structures a été la publication de la tranche d'Univers du CfA-II (de Lapparent *et al.*, 1986, voir figure 1.2). On y voit clairement déjà la présence de structures avec des régions vides et d'autres très denses.

Depuis, de nombreux sondages ont suivi. Le but du *Southern Sky Redshift Survey* (SSRS) (da Costa *et al.*, 1991) était de compléter le CfA original en cartographiant les galaxies dans l'hémisphère sud. L'*Optical Redshift Survey* (ORS) (Santiago *et al.*, 1995), d'une profondeur équivalente à celle du CfA ($80 h^{-1} \text{ Mpc}$, magnitude limite en B de 14.5) a lui couvert les deux tiers de la surface du ciel et collecté quelques ~ 8500 *redshifts*. Ce

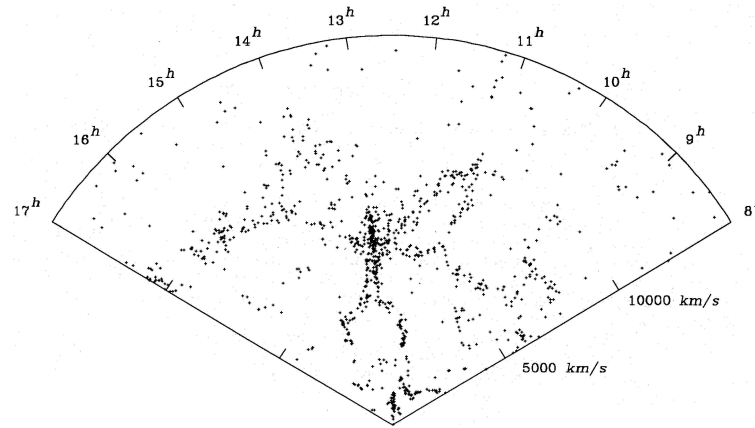


FIG. 1.2 – “Slice of the Universe”, de Lapparent *et al.* (1986)

sondage a permis entre autres une étude du champ de densité local, de la fonction de luminosité, de la fonction de corrélation et des dispersions de vitesses des galaxies.

Le *Stromlo-Automated Plate Machine* (Loveday *et al.*, 1996) contient 1797 galaxies avec une magnitude $b_J < 17.15$ sélectionnées aléatoirement, avec un échantillonnage spatial de 1 pour 20, à partir des images du sondage APM. Ce dernier couvre une surface de 4300 degrés carrés. Le *redshift* moyen est de ~ 0.05 . La profondeur du *Durham/UK Schmidt Telescope redshift survey* (Durham/UKST) (Ratcliffe *et al.*, 1998a) lui est similaire mais l'échantillonnage spatial est de 1 pour 3. La plus grande vitesse mesurée est de $\sim 35000 \text{ km.s}^{-1}$.

Les données du satellite IRAS (*Infrared Astronomical Satellite*) ont été utilisées pour réaliser des sondages spectroscopiques de l'Univers proche observé dans l'infrarouge. Ainsi, on retiendra les sondages à 2-Jy² de Strauss *et al.* (1992), à 1.2-Jy de Fisher *et al.* (1995), le QDOT (*Queen Mary and Westfield college, Durham, Oxford and Toronto*) de Lawrence *et al.* (1999) et le sondage PCSz (*Point Source Catalog z*) de Saunders *et al.* (2000). Ce dernier contient quasiment toutes les sources détectées avec IRAS au flux limite de 0.6-Jy.

De retour dans l'optique, le sondage ESP (*European Southern Observatory Slice Project*; Vettolani *et al.* (1998)) consiste en une tranche du ciel de 1×22 degrés carrés pour une profondeur de $600 \text{ h}^{-1} \text{ Mpc}$. L'une des questions posées au travers de ce sondage fut la nature fractale de la distribution des galaxies à grande échelle. Enfin, le *Las Campanas Redshift Survey* (LCRS; Shectman *et al.* (1996)) a une géométrie similaire avec 6 bandes parallèles de 1.5×90 degrés carrés pour une profondeur de $750 \text{ h}^{-1} \text{ Mpc}$, ce qui correspond à un décalage spectral $z \sim 0.25$. Près de 24000 *redshifts* ont été mesurés. Ce fut le premier sondage profond, avec un assez grand volume pour étudier correctement les propriétés de l'Univers local.

Afin de décrire *totale*ment l'Univers local, deux grand projets de sondages spectroscopiques profonds ont été lancés : le *Two-Degree Field galaxy redshift survey* (2dFGRS; Colless *et al.* (2001)) et le *Sloan Digital Sky Survey* (SDSS; Abazajian *et al.* (2004))

Le 2dFGRS est un projet anglo-australien qui utilise un spectrographe multi-objet sur le télescope de 3.9 mètres de l'AAO (*Anglo-Australian Observatory*). Il est capable d'enregis-

²1 Jy(jansky) = $10^{-26} \text{ W.m}^{-2} \text{ Hz}^{-1}$

trer le spectre de 400 objets simultanément dans un champ de vue circulaire de 2 degrés au moyen de fibres optiques positionnées automatiquement. Les cibles sont sélectionnées dans le catalogue APM jusqu'à une magnitude $b_J = 19.45$. La surface totale est de 2000 degrés carrés, pour ~ 230000 *redshifts*. La figure 1.3 montre la distribution de ces galaxies en fonction de leur coordonnée R.A. (ascension droite) et *redshift*. Les structures (régions denses, vides) sont clairement visibles. La figure 1.4 montre par exemple deux structures importantes présentes dans les données du 2dFGRS. Le relevé du 2dF est aujourd'hui terminé.

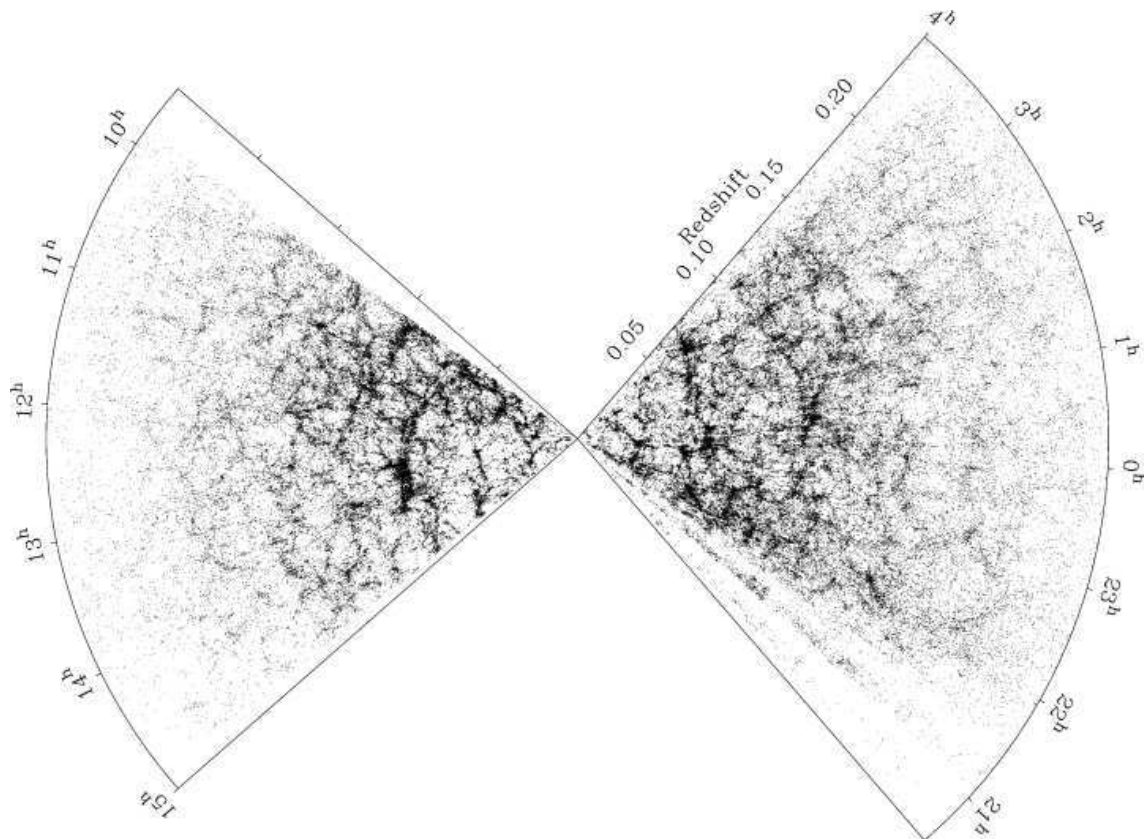


FIG. 1.3 – Projection dans l'espace (ascension droite, *redshift*) des galaxies du 2dFGRS (Colless *et al.*, 2001, figure 18).

Le SDSS, par contre, est toujours en cours d'acquisition. Près de deux cent personnes sont impliquées dans ce sondage au travers de 13 instituts américains, européens et japonais. L'objectif est d'observer l'équivalent de la moitié du ciel dans plusieurs filtres u (3550Å), g (4770Å), r (6230Å), i (7620Å) et z (8870Å), et ensuite de mesurer le *redshift* de $\sim 10^6$ galaxies.

La fonction de sélection est la principale différence entre le 2dFGRS et le SDSS, mis à part le volume de données et la couverture du ciel. En effet, la sélection des objets du 2dFGRS est faite dans le bleu alors que celle du SDSS est faite dans le rouge à $r_{\text{lim}} = 17.77$.

D'un point de vue quantitatif, outre les résultats sur l'agrégation présentés dans la partie 1.5, les distributions en *redshifts* observées dans l'Univers proche sont maintenant bien connues. Celle du 2dFGRS est par exemple représentée figure 1.5. On distingue ainsi

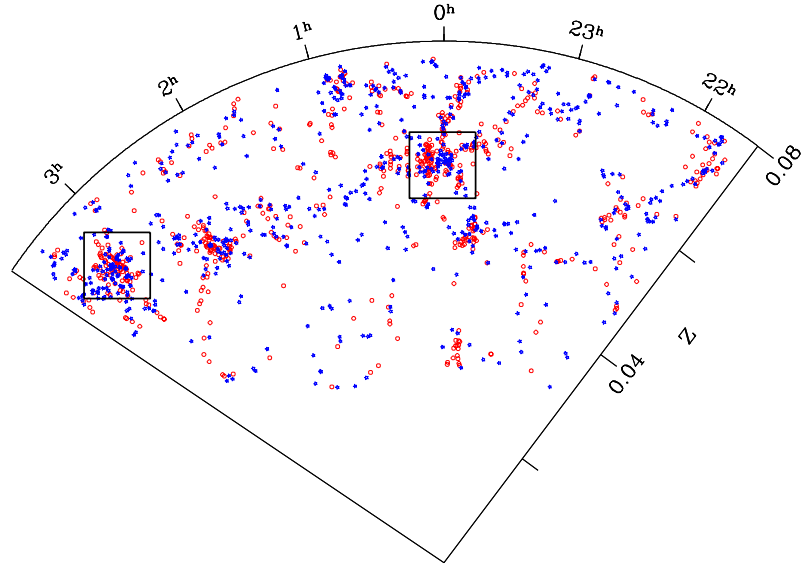


FIG. 1.4 – Distribution spatiale de galaxies d’une partie du champ sud du 2dFGRS avec $-19 \leq M_{b_j} - 5\log(h) \leq -18$. Les étoiles bleues représentent les galaxies de types tardifs et les cercles rouges les galaxies précoces. Les deux boîtes indiquent les positions de structures importantes. (Norberg *et al.*, 2002, figure 3)

deux pics principaux à $z \sim 0.09$ et $z \sim 0.11$.

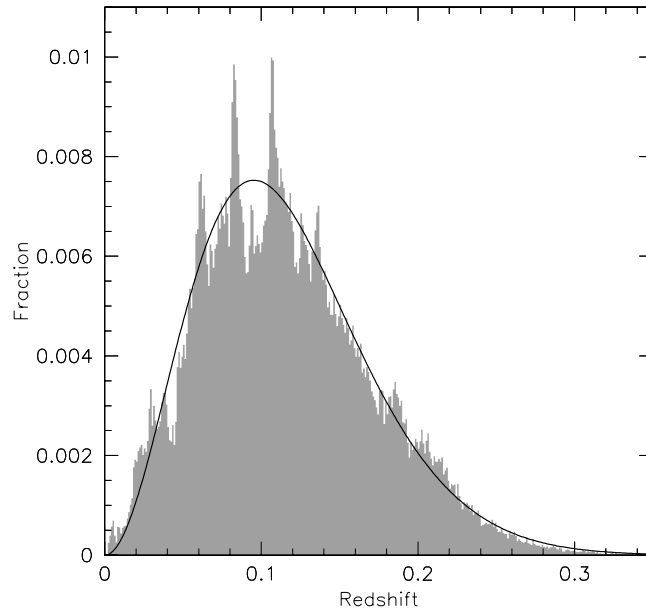


FIG. 1.5 – Distribution en *redshifts* du 2dFGRS avec, superposé, un simple lissage analytique (Colless *et al.*, 2001, figure 17)

1.3.4 Structures dans l'Univers lointain ($z \gtrsim 0.3$)

D'un point de vue qualitatif, l'Univers lointain présente lui aussi des structures. Mis en évidence par [Le Fèvre et al. \(1996\)](#) dans le cadre du sondage spectroscopique CFRS, des pics de densité sont clairement présents jusqu'à des *redshifts* de $z \sim 1$ (figure 1.6). On observe une alternance de régions vides et riches en galaxies.

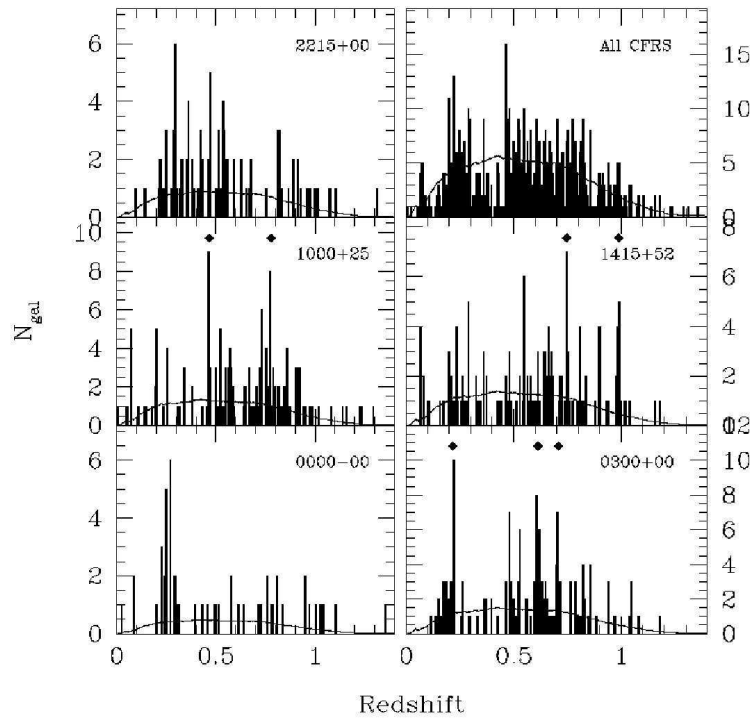


FIG. 1.6 – Distribution en *redshifts* des champs du CFRS ([Le Fèvre et al., 1996](#)). On observe une alternance de pics dans la distribution des galaxies.

À de plus grand *redshifts*, les observations de [Steidel et al. \(1998\)](#) permettent de mettre en évidence de grandes structures de galaxies à $z \sim 3$. Les galaxies sont identifiées en photométrie (voir partie 1.4.2.4) et leurs décalages spectraux sont ensuite mesurés spectroscopiquement avec le spectrographe LRIS sur le télescope Keck dont le diamètre de 10 mètres permet l'observation de tels objets. Une surdensité importante d'objets est ainsi détectée à $\langle z \rangle = 3.090$. [Adelberger et al. \(1998\)](#) rapportent la distribution en *redshifts* de tels objets dans six champs différents (figure 1.7). La variation de champ à champ de ces distributions est principalement due à l'incomplétude spectroscopique d'après leurs auteurs mais elles mettent en évidence l'alternance de régions denses et vides à une époque où l'Univers était âgé de $\sim 20\%$ de son âge actuel.

Les sondages de galaxies alliant photométrie multi-couleur et spectroscopie ont ainsi mis clairement en évidence l'existence de structures à différents *redshifts*, de l'Univers local jusqu'à l'Univers lointain. Cette description qualitative doit maintenant être quantitative, ceci afin de suivre l'évolution et la formation de ces structures. La partie suivante se destine donc à faire un tour d'horizon des principales propriétés d'agréations des galaxies dans l'Univers.

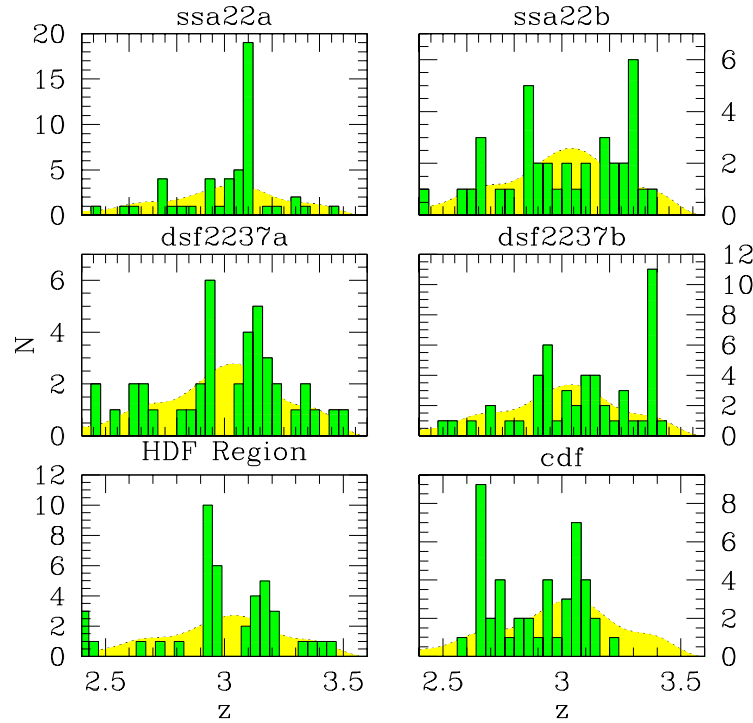


FIG. 1.7 – Distribution en *redshift* de 6 champs où des objets à $z \sim 3$ (LBGs, partie 1.4.2.4) sont recherchés et identifiés (Adelberger *et al.*, 1998, figure 2).

1.4 La fonction de corrélation à 2 points : un outil majeur

1.4.1 Un outil pour quantifier la distribution des galaxies et de la matière : définition et estimateur

Totsuji et Kihara (1969) sont les premiers à calculer la fonction de corrélation spatiale à deux points $\xi(r)$. Cette fonction représente l'excès de probabilité qu'une galaxie ait un voisin situé à une distance r , comparée à une distribution aléatoire poissonnienne. La probabilité de trouver une galaxie dans un élément de volume δV et de densité moyenne ρ_0 étant $\rho_0 \delta V$, la probabilité de trouver une paire de galaxies avec une galaxie dans le volume δV_1 et l'autre dans le volume δV_2 est donc

$$\delta P = \rho_0^2 (1 + \xi(r_{12})) \delta V_1 \delta V_2 \quad (1.5)$$

avec r_{12} la distance séparant les deux galaxies considérées. De façon consistante avec le principe d'homogénéité et d'isotropie de l'Univers, cette probabilité ne dépend que de la distance r_{12} et non de la position. Dans le cas d'une distribution aléatoire, les probabilités δP_1 et δP_2 sont indépendantes. La probabilité jointe est alors égale au produit des deux. L'équation 1.5 devient alors

$$\delta P = \rho_0^2 \delta V_1 \delta V_2 \quad (1.6)$$

La valeur de la fonction de corrélation est alors nulle quelque soit la distance r_{12} . Si, pour une échelle de distance donnée, la valeur de $\xi(r)$ est positive, alors la distribution présente une région de surdensité. Si elle est négative, il y a une sous-densité.

Une autre approche revient à considérer la fonction de corrélation comme la fonction d'auto-corrélation du contraste du champ de densité $\delta(\vec{x})$ défini par

$$\delta(\vec{x}) = \frac{\rho(\vec{x}) - \langle \rho \rangle}{\langle \rho \rangle} \quad (1.7)$$

Elle prend alors la forme de l'équation 1.8

$$\xi(\vec{r}) = \langle \delta(\vec{x}) \delta(\vec{x} + \vec{r}) \rangle \quad (1.8)$$

Les premières mesures de la fonction de corrélation par [Groth et Peebles \(1977\)](#) et [Davis et Peebles \(1983\)](#) mettent en évidence sa forme caractéristique en loi de puissance, pour des échelles inférieures à $20 h^{-1} \text{Mpc}$:

$$\xi(r) = \left(\frac{r}{r_0} \right)^{-\gamma} \quad (1.9)$$

Des études récentes ([Guzzo et al., 1991](#); [Gaztañaga et Juszkiewicz, 2001](#)) semblent indiquer que cette loi de puissance n'est pas universelle. Un changement de la pente γ est observé à des échelles de $1-2 h^{-1} \text{Mpc}$. Il est expliqué par un changement de régime : à petite échelle, la fonction de corrélation est dominée par le nombre de paires de galaxies situées dans le même halo de matière noire et à grande échelle par les séparations entre galaxies de halos différents. Le modèle de halo HOD (page 47) arrive à reproduire ces observations ([Zehavi et al., 2004, 2005b](#); [Phleps et al., 2005](#)).

On appelle r_0 longueur de corrélation. C'est une longueur caractéristique qui quantifie l'état d'agrégation d'un échantillon de galaxies. La probabilité qu'une galaxie ait une voisine située à la distance $r=r_0$ est deux fois plus grande que dans le cas d'une distribution aléatoire poissonnienne ($\xi(r) = 1$). Ou encore, si une population de galaxie est plus concentrée qu'une autre, alors, pour une échelle de distance donnée, la probabilité pour qu'une galaxie de cette population ait une voisine sera plus importante, et donc la valeur de r_0 aussi. Enfin, pour donner une image, considérant deux populations de galaxies de longueurs de corrélation identiques, celle dont la pente γ est la plus importante présentera une structure plus "allongée" traduisant ainsi la diminution du nombre de paires de galaxies de grandes séparations au profit de celle de petites séparations.

Afin d'obtenir ces mesures de $\xi(r)$, différents estimateurs ont été développés. Ils sont basés sur la définition plus pratique qui consiste à faire le rapport entre le nombre $N_g(r)$ de couple de galaxies séparées d'une distance r et $N_r(r)$ celui d'une distribution aléatoire poissonnienne :

$$1 + \xi(r) = \frac{N_g(r)}{N_r(r)} \quad (1.10)$$

Le nombre de galaxies distribuées aléatoirement étant dans la pratique généralement supérieur d'un facteur 10 au moins au nombre de galaxies étudiées, un facteur de normalisation vient ajuster le terme de droite de l'équation 1.10. Pour calculer $\xi(r)$, il s'agit donc de réaliser un comptage de paires, pour toutes les échelles de distance. Le premier estimateur utilisé fut celui de [Davis et Peebles \(1983\)](#) :

$$\xi(r) = \frac{DD(r)}{DR(r)} - 1 \quad (1.11)$$

$DD(r)$ représente le nombre de paires de galaxies séparées d'une distance r , $DR(r)$ représente celui entre une galaxie et un point du catalogue généré aléatoirement.

Une autre solution a été proposée par [Hamilton \(1993\)](#) faisant intervenir le calcul du nombre de paires $RR(r)$ interne au catalogue de points aléatoires :

$$\xi(r) = \frac{DD(r) \times RR(r)}{DR(r)^2} - 1 \quad (1.12)$$

Mais nous ne considérerons par la suite qu'uniquement l'estimateur de [Landy et Szalay \(1993\)](#) qui est le plus efficace car il minimise la variance :

$$\xi(r) = \frac{DD(r) - 2DR(r) + RR(r)}{RR(r)} \quad (1.13)$$

L'estimateur classique de [Landy et Szalay \(1993\)](#) convient bien à la taille des échantillons considérés dans cette thèse. Pour des échantillons de tailles encore plus importantes, comme par exemple l'utilisation d'échantillons de *redshifts* photométriques, il sera intéressant de développer un outil de type KD-tree ([Gray et al., 2004](#); [Zhang et Pen, 2005](#)) pour rendre le temps de calcul acceptable.

1.4.2 Fonction de sélection et comparaison des mesures

Le choix de l'échantillon de galaxies à étudier avec la fonction de corrélation à deux points est important. En effet, toute l'interprétation des résultats dépend de la connaissance des propriétés des objets observés et donc de la fonction de sélection. Les mesures obtenues pour deux sondages différents ne pourront être comparées directement que si les critères de sélection des galaxies sont effectivement comparables. On décrit dans les sous sections suivantes les notions de "population globale", "types de galaxies", "galaxies LBG" et "galaxie ERO".

La comparaison de résultats obtenus indépendamment, avec la même fonction de sélection, doit aussi prendre en compte la variance cosmique. Ceci est surtout valable dans le cas où les champs couverts sont de petites tailles. En effet, si le champ A compte par exemple une région très dense en son sein (un amas de galaxie par exemple), la mesure de la longueur de corrélation sera plus importante que celle mesurée dans le champ B, où les galaxies sont plus distribuées. Pour de grands champs, cet effet est moyenné, mais il est aussi envisageable d'observer plusieurs champs indépendants. Dans tous les cas, il est important de réaliser de grands sondages et/ou d'être capable de bien quantifier la variance cosmique d'une manière ou d'une autre. [Somerville et al. \(2004\)](#) estiment que pour des champs de la taille du champ CDF-S du sondage GOODS ($10' \times 16'$, pour un volume de $\sim 10^5 \text{ Mpc}^3$), la variance attendue pour des galaxies très fortement agrégées de type EROs (section 1.4.2.5) est de $\sim 40\text{-}60\%$. Elle est moins importante pour des populations moins agrégées : $\sim 10\text{-}20\%$ pour des galaxies LBGs (voir section 1.4.2.4). Et comme attendu, la variance cosmique diminue très rapidement dès que le volume sondé augmente. La variance de la population LBG passe ainsi à moins de $\sim 10\%$ pour un volume d'Univers supérieur à $\sim 10^6 \text{ Mpc}^3$.

1.4.2.1 Population globale

Le critère de sélection le plus simple d'un échantillon de galaxies revient à considérer tous les objets plus brillants qu'une certaine magnitude apparente. Aucun autre critère n'étant introduit, la population de galaxies ainsi obtenue n'a pas de propriété particulière. On parle alors de population globale.

Une limite en magnitude apparente se traduit, à un *redshift* donné, par une population plus brillante qu'elle ne l'est réellement intrinsèquement. Les galaxies les plus faibles ne sont en effet pas observées (biais de Malmquist). Il n'en reste pas moins que l'échantillon obtenu est globalement représentatif de l'ensemble des galaxies situées à ce décalage spectral. Son analyse fournit ainsi des propriétés générales sur les galaxies à un *redshift* donné. Malheureusement, la comparaison de résultats obtenus à des *redshifts* différents n'est alors pas directe puisque qu'une limite en magnitude apparente se traduit par une limite en magnitude absolue qui devient de plus en plus brillante quand le *redshift* augmente.

1.4.2.2 Sélection en fonction de la luminosité absolue

On peut vouloir étudier l'agrégation des galaxies d'un échantillon en fonction de leurs luminosités intrinsèques (absolues). Il faut alors définir des échantillons selon leurs magnitudes absolues mesurées dans une certaine bande. Le calcul de la magnitude absolue d'une galaxie nécessite de connaître son décalage spectral ainsi que les paramètres cosmologiques décrivant le modèle d'Univers. Seul un sondage spectroscopique permet donc d'étudier l'influence de la luminosité.

On appelle un échantillon limité en volume un échantillon dans lequel seules des galaxies plus brillantes qu'une magnitude absolue fixée sont sélectionnées.

Dans un sondage sélectionné en magnitude apparente, les galaxies les plus éloignées sont statistiquement plus brillantes (biais de Malmquist, figure 1.8). Un échantillon limité en volume permet donc de suivre l'évolution d'un paramètre physique (longueur de corrélation par exemple) en fonction du *redshift* pour une même classe d'objets, plus brillants que la magnitude absolue mesurée au plus grand décalage spectral. On a ainsi représenté sur la figure 1.8 une boîte qui contient uniquement des objets plus brillants que $M_{AB} \leq -18 + 5\log(h)$, dans l'intervalle en *redshift* $z=[0.5-0.8]$. Au-delà de $z \sim 0.8$, les objets les plus faibles que la limite choisie ne sont rapidement plus visibles.

1.4.2.3 Sélection en fonction du type

type morphologique

C'est Hubble (1936) qui introduit la première classification morphologique des galaxies (figure 1.9), mise à jour par de Vaucouleurs (1974).

Les premières études de l'agrégation des galaxies en fonction de la morphologie nécessitaient de classer les galaxies à l'oeil. C'est par exemple le cas des galaxies du catalogue Uppsala, où la classification a été réalisée à partir de l'analyse de plaques photographiques (Davis et Geller, 1976).

Une classification à l'oeil devient difficile dès lors que les objets sont faibles, peu lumineux. Elle est alors sujette à caution. La comparaison d'études indépendantes basées

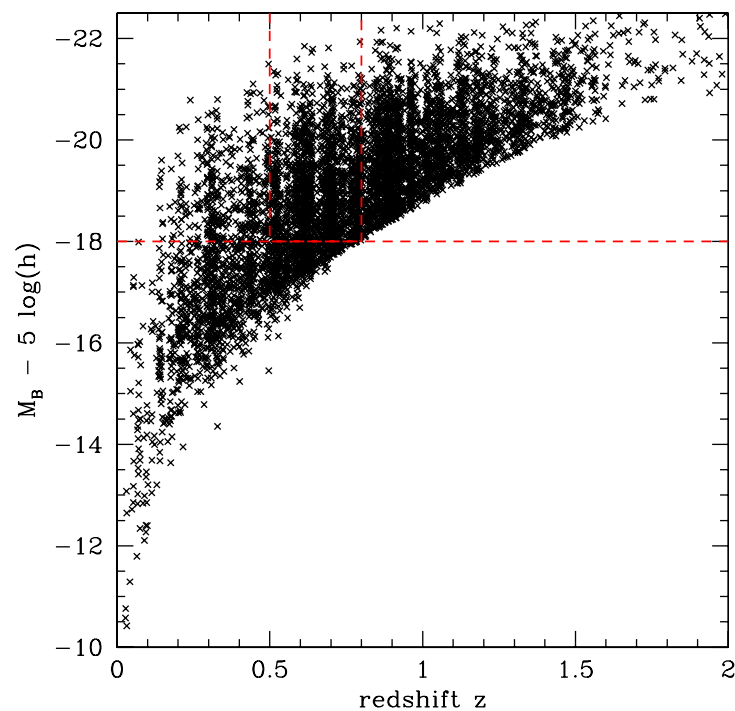


FIG. 1.8 – Magnitude absolue en bande B des galaxies du sondage VVDS en fonction du *redshift*. La ligne en pointillé rouge représente une coupure à $M_{AB} \leq -18 + 5\log(h)$ d'un échantillon limité en volume. Celui-ci est complet jusqu'à $z \sim 0.8$.

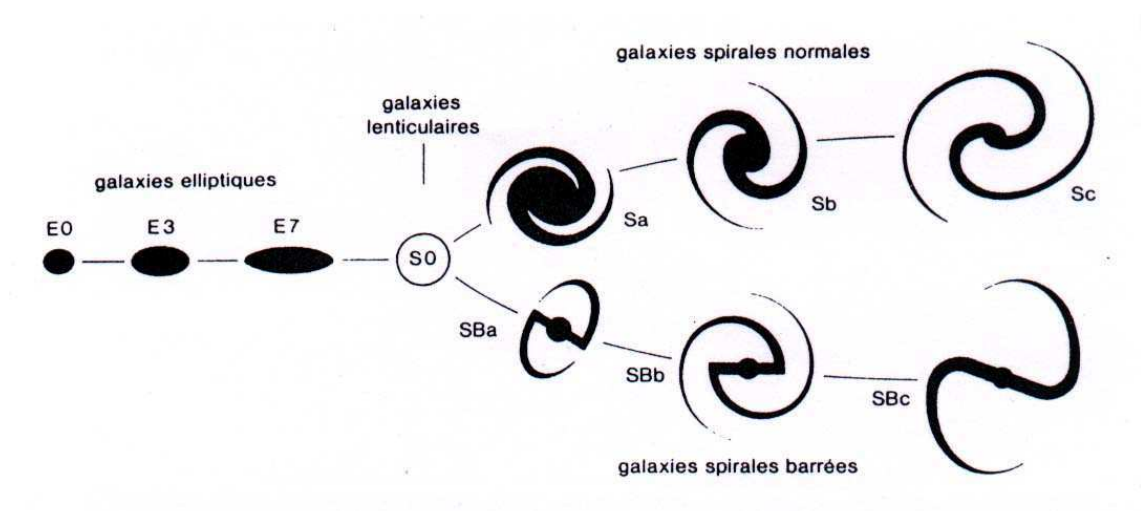


FIG. 1.9 – Classification morphologique de Hubble

chacune sur ce genre de classification peut alors montrer des différences.

Une classification plus objective selon des critères morphologiques consiste à quantifier la concentration et l'asymétrie des galaxies à partir de programmes informatiques automatisés tels que GALFIT (Peng *et al.*, 2002) ou GIM2D (Simard *et al.*, 2002). Par exemple, dans le cadre du SDSS, Zehavi *et al.* (2002) divisent leur échantillon en deux selon un critère de concentration, rapport entre les rayons isophotales qui contiennent respectivement 90 et 50% du flux de la galaxie considérée. Cette classification permet de distinguer les galaxies précoces dominées par leur bulbe et les galaxies tardives dominées par leur disque. À grands *redshifts* $z \geq 0.6$, seules les images obtenues à partir du télescope spatial *Hubble* permettent d'obtenir une information morphologique quantitative avec la mesure des composantes "bulbe" et "disque" et leurs grandeurs caractéristiques (rayon, brillance). De telles classifications mettent en évidence une forte évolution de la distribution en types avec le *redshift* (Brinchmann *et al.*, 1998).

couleur/type spectral

Une autre méthode d'analyse de l'agrégation des galaxies consiste à étudier celle-ci en fonction de critères de couleur. Cette approche permet de réaliser une classification objective de galaxies faibles et/ou situées à grands *redshifts*.

Par exemple, à partir d'images obtenues avec le satellite Hubble dans le cadre du sondage *Medium Deep Survey*, Neuschaefer *et al.* (1995) étudient la fonction de corrélation angulaire en fonction de la couleur (V-I). Landy *et al.* (1996) séparent leurs deux échantillons de ~ 2900 galaxies en fonction de leurs couleurs mesurées en (U-B), (U-R) et (B-R). Roche *et al.* (1996) obtiennent des images profondes jusqu'à $B \sim 25.5$ et $R \sim 24.5$ avec le télescope Isaac Newton à La Palma. La qualité de ces données leur permet de diviser leur échantillon de ~ 7000 galaxies selon que leurs couleurs (B-R) soit supérieures ou inférieures à 1.5.

Plus récemment, une analyse en couleur a été menée dans le cadre du SSRS par Willmer *et al.* (1998). Le SSRS étant un sondage spectroscopique, une classification selon la couleur au repos est rendue possible. Ainsi, la coupure en couleur choisie est $(B - R)_0 = 1.3$ qui correspond grossièrement à une galaxie de type S_{bc} . Ces auteurs ne considèrent que les galaxies dont la distance n'excède pas 8000 km.s^{-1} et définissent ainsi 387 galaxies bleues et 219 rouges.

Une simple coupure en couleur permet généralement de définir un échantillon bleu dominé par des galaxies de type tardif et un échantillon rouge dominé par des galaxies précoces. Mais la couleur observée d'un type de galaxie donné étant une fonction du *redshift*, Brown *et al.* (2000) préfèrent définir un critère de sélection selon deux couleurs (U-B) et (B-R) pour permettre une sélection de type de galaxies sur une large plage en *redshifts* sans trop dépendre des paramètres cosmologiques. La population bleue doit ainsi vérifier

$$(B - R) \leq 1 \text{ ou } \begin{cases} (B - R) \leq 1.9 \\ \text{et} \\ (U - B) \leq 1.4 \end{cases} \quad (1.14)$$

alors que la population rouge est définie par

$$\left\{ \begin{array}{l} (B - R) \geq 1.9 \\ \text{et} \\ (U - B) \geq -0.6 \end{array} \right. \quad \text{ou} \quad \left\{ \begin{array}{l} (B - R) \geq 1 \\ \text{et} \\ (U - B) \geq 0.4 \end{array} \right. \quad (1.15)$$

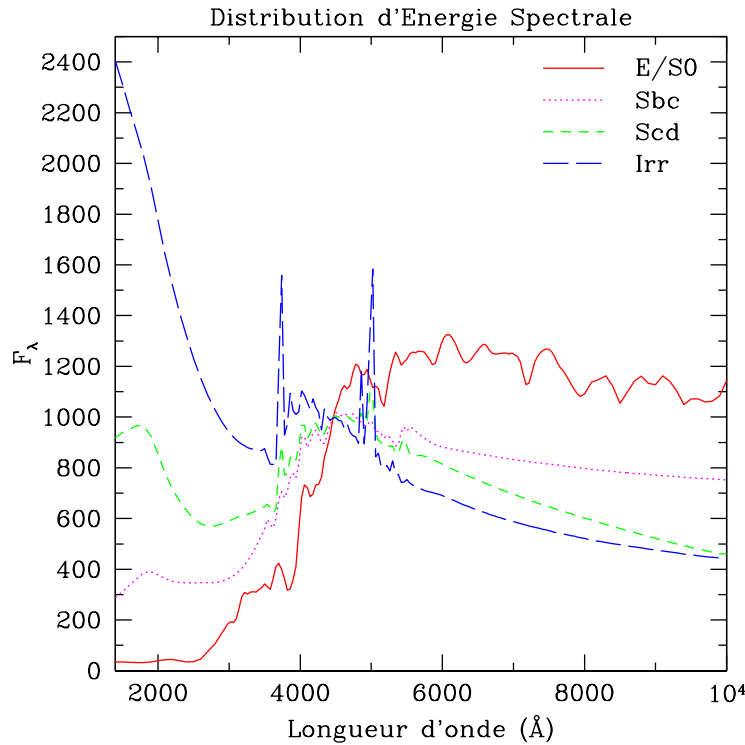
Dans le cadre du SDSS, [Strateva et al. \(2001\)](#) exhibent une distribution en couleur $u^* - r^*$ bimodale. Ils montrent que la population bleue contient principalement des galaxies de type morphologique tardif (spirales), alors que la classe rouge est composée en majorité de galaxies dominées par leur bulbe. [Zehavi et al. \(2002\)](#) utilisent alors cette bimodalité pour séparer en deux ~ 29300 galaxies autour de la couleur au repos $u^* - r^* = 1.8$. La classification ainsi définie conduit à une répartition rouge-bleue de $\sim 2/3 - \sim 1/3$.

Complétant cette première étude avec cette fois ~ 200000 galaxies du SDSS, [Zehavi et al. \(2005b\)](#) étudient la dépendance à la couleur de l'agrégation des galaxies, critère le mieux estimé dans ces données. La division "rouge/bleue" se fait autour de la couleur $(g - r) = 0.7$, avec une dépendance linéaire à la magnitude absolue M_r qui permet d'avoir à peu près autant de galaxies bleues que rouges pour chaque intervalle de luminosité. Les galaxies rouges restent majoritaires parmi les objets brillants, et inversement à magnitudes faibles où les galaxies bleues dominent.

Une version plus raffinée de la classification de galaxies en fonction d'une couleur $((V-I), (B-R), \dots)$ consiste à utiliser l'information spectrale et multi-longueur d'onde. Des classifications qui se veulent objectives ont ainsi été utilisées depuis peu. Ainsi, [Loveday et al. \(1999\)](#) basent leur classification sur la largeur équivalente au repos des raies de l'oxygène [OII] et de l'hydrogène $H\alpha$. Les objets présentant des raies en émission sont associés à des types tardifs en train de former des étoiles.

Dans le cadre du 2dFGRS, [Norberg et al. \(2002\)](#) utilisent une classification basée sur une décomposition de chaque spectre de galaxie en composantes principales (*principal-component analysis*, PCA). Cette classification permet de décrire chaque spectre selon un seul paramètre η qui varie de façon continue ([Madgwick et al., 2002](#)). La valeur de η est fonction de la largeur des raies d'émission du spectre considéré au repos et peut être interprété comme un taux de formation d'étoiles de chaque galaxie. Cette classification présente aussi la propriété d'être corrélée avec une classification morphologique réalisée de façon subjective à partir d'image en bande b_J ([Madgwick et al., 2002](#)). [Norberg et al. \(2002\)](#) définissent ainsi deux classes de galaxies selon cette méthode : galaxies précoces et galaxies tardives.

Une dernière méthode utilisée pour attribuer un type à une galaxie est d'utiliser son spectre et ses magnitudes apparentes mesurées dans différents filtres pour lui associer un objet de référence dont les propriétés spectro-photométriques sont connues. Ainsi, [Budavári et al. \(2003\)](#) se basent sur les quatre spectres observés par [Coleman et al. \(1980\)](#) et représentés figure 1.10. Ces spectres sont considérés comme représentatifs de galaxies elliptiques, spirales et irrégulières de l'Univers proche et sont utilisés pour classer au mieux de grands catalogues de galaxies. [Budavári et al. \(2003\)](#) réalisent en plus une extrapolation de ces quatre spectres afin de trouver le meilleur modèle pour chacune des 2 millions de galaxies issues du SDSS. Les types sont attribués au travers d'un code de calcul de *redshifts* photométriques. Cette classification s'apparente à une classification en couleur avec des galaxies précoces rouges et des galaxies tardives bleues.

FIG. 1.10 – Spectres de références (Coleman *et al.*, 1980)

1.4.2.4 Galaxies à discontinuité Lyman (LBG)

L'étude de galaxies situées à de grands *redshifts* fait appel à différentes techniques pour identifier ces populations. La plus connue s'appuie sur la détection de la cassure à 912Å au repos du spectre des galaxies, la discontinuité de Lyman. Cette variation brutale du spectre est due à l'ionisation totale de l'hydrogène depuis son état fondamental par des photons plus énergétiques que 13.6 eV. Cette discontinuité a été observée dans l'ultra-violet et est particulièrement marquée pour les galaxies à sursauts de formation d'étoiles. Lorsque ces galaxies sont situées à un certain *redshift*, cette discontinuité est décalée à de plus grandes longueurs d'ondes et apparaît dans le visible pour $z \geq 3$ et $\lambda \geq 4000\text{Å}$. L'utilisation de plusieurs filtres optiques permet de détecter cette discontinuité. Un objet détecté en BVRI mais pas en U sera potentiellement situé à $z \sim 3$, alors qu'un objet pas ou peu détecté en U et B mais clairement identifié en VRI pourra se localiser à $z \sim 4$. En pratique, on isole de tels candidats à grand décalage spectral par une sélection dans un diagramme couleur-couleur. Par exemple, une galaxies à $z \sim 3$ sera rouge en (U-B) (car pas détectée en U) et bleue en (B-I). Une confirmation spectroscopique du *redshift* reste tout de même nécessaire.

Forts du succès de cette technique appliquée initialement pour la recherche de quasars, Steidel *et al.* (1996), après avoir identifié une fonction de sélection (équation 1.16) à partir de filtres U_n , G, \mathcal{R} et I (figure 1.11), confirment 16 LBGs (*Lyman-Break Galaxies*) dans l'intervalle $3.0 \leq z \leq 3.5$ sur 23 candidats. La luminosité de ces objets ainsi que leur densité surfacique ($\sim 0.08 [\text{minutes d'arc}]^{-2}$), étant très faibles, des images profondes sur de *grands* champs sont requises.

$$\begin{cases} \mathcal{R} \leq 25.5 \\ G - \mathcal{R} \leq 1.2 \\ U_n - G \geq G - \mathcal{R} + 1.5 \end{cases} \quad (1.16)$$

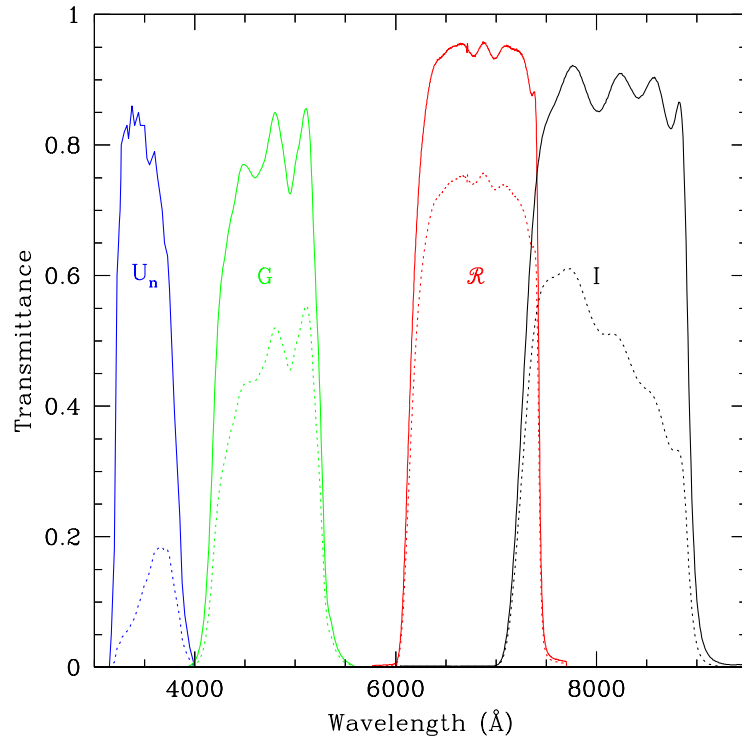


FIG. 1.11 — Courbe de transmission des filtres U_n , G , \mathcal{R} et I (Steidel *et al.*, 2003, figure 1). En pointillé, la performance effective prenant en compte l'efficacité des capteurs CCD et l'absorption atmosphérique.

1.4.2.5 Galaxies EROs

Alors que les sondages optiques observent l'Univers jusqu'à $z \sim 1$ et que les techniques *Lyman-Break* sélectionnent des objets au-delà de $z \sim 2.5$, la réalisation de sondage de galaxies sélectionnées dans le proche infrarouge a permis une première exploration de la population de galaxies de type précoce pour $1.3 \leq z \leq 2.5$ (Cimatti *et al.*, 2002, sondage K20). Les galaxies massives et rouges deviennent faibles en optique à partir de $z \sim 1$ mais restent relativement brillantes observées à de plus grandes longueurs d'onde. De plus la luminosité reçue en (proche) infrarouge suit d'avantage la masse totale stellaire et est moins affectée par l'absorption par les poussières ou par les étoiles jeunes très brillantes. Les mesures de masse totale stellaire sont alors plus facilement comparables aux modèles que lorsqu'une sélection dans le visible est effectuée.

Au cours des années 1990, les premiers sondages de galaxies dans le proche infrarouges ont conduit à la découverte de galaxies dites extrêmement rouges (EROs pour

Extremely Red Objects) avec des couleurs typiques de $(R-K) \geq 5$ ou 6. Les modèles indiquent que ces couleurs sont typiques de galaxies elliptiques à $z \sim 1$ ou plus, de galaxies à grands *redshifts*, poussiéreuses, en train de former des étoiles, de galaxies à noyaux actifs obscurcis et d'étoiles de faibles masses. Des observations laissent à penser que 70 à 80% de ces objets sont des galaxies elliptiques évoluées et que les plus rouges des autres correspondent à des galaxies poussiéreuses à formation stellaire à $z > 1$.

1.4.3 Fonction de corrélation et biais

Nous avons introduit la notion de distribution biaisée des galaxies dans la section 1.2.2. [Kaiser \(1984\)](#) traduit en terme de fonction de corrélation sa théorie des pics et obtient, en étudiant les amas d'Abell, la relation

$$A \equiv \frac{\xi_{\text{am}}}{\xi_{\text{m}}} \quad (1.17)$$

avec \sqrt{A} le biais introduit, ξ_{am} la fonction de corrélation spatiale des amas et ξ_{m} celle des pics de densité sous-jacente.

Avec le calcul des premiers modèles à N corps par [Davis et al. \(1985\)](#), dans le cadre d'un Univers CDM, des études sur le développement des structures non-linéaires purent être menées. Elles conduisirent à la conclusion que la fonction de corrélation des galaxies devait être une version amplifiée de celle de la masse. On parle alors de formation biaisée des galaxies où les galaxies brillantes, massives, se forment au niveau des maximums locaux du champ de densité, si celui-ci dépasse un certain seuil (figure 1.12).

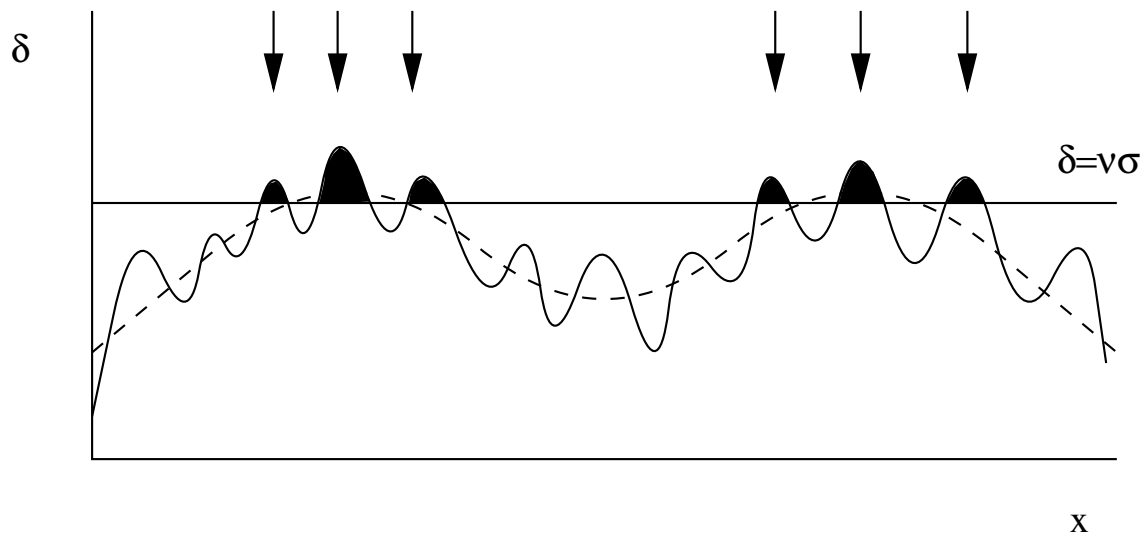


FIG. 1.12 – Schéma des fluctuations d'un champ gaussien aléatoire $\delta(x)$ ([Peacock, 1999](#), figure 8). Les flèches indiquent les maximums locaux qui dépassent un certain seuil et où vont se former les galaxies.

Le biais linéaire b est souvent décrit comme un facteur de proportionnalité indépendant

de l'échelle considérée. Il relie les fluctuations de densité de la matière et des galaxies par :

$$\left(\frac{\delta\rho}{\rho}\right)_{\text{galaxies}} = b \left(\frac{\delta\rho}{\rho}\right)_{\text{masse}} \quad (1.18)$$

Exprimé en terme de fonction de corrélation, ce biais est dépendant de l'échelle.

Une autre définition relativement équivalente et souvent utilisée est celle qui décrit la variance d'un comptage de galaxies à l'intérieur d'une sphère de $8 h^{-1} \text{Mpc}$ de rayon par rapport à la variance de la distribution de masse sous-jacente.

$$b = \frac{\sigma_8(\text{galaxies})}{\sigma_8(\text{masse})} \quad (1.19)$$

où σ^2 est l'intégrale volumique de la fonction de corrélation spatiale $\xi(r)$

$$\sigma^2(r) = \frac{1}{V^2} \int_V \xi(r) d^3r_1 d^3r_2 \quad (1.20)$$

Si $\xi(r)$ est une loi de puissance, σ_8 peut encore s'exprimer sous la forme ([Peebles, 1980](#))

$$\sigma_8 = \sqrt{C_\gamma \left(\frac{r_0}{8 h^{-1} \text{Mpc}} \right)^\gamma} \quad (1.21)$$

avec

$$C_\gamma = \frac{72}{(3-\gamma)(4-\gamma)(6-\gamma)2^\gamma} \quad (1.22)$$

On retiendra que la mesure du biais peut être contrainte par la mesure de la fonction de corrélation d'objets de natures différentes.

1.5 État des lieux : mesures de l'agrégation des galaxies

Le but de cette section est de réaliser une compilation aussi exhaustive que possible des mesures d'agrégation des galaxies les plus significatives effectuées ces dernières années, depuis l'Univers proche, avec différentes fonctions de sélection (population globale, amas, luminosité, morphologie, couleur - voir section [1.5.1](#)) jusqu'aux galaxies les plus distantes (section [1.5.2](#)).

Ces mesures constituent une référence à laquelle je comparerai mes nouvelles mesures de l'Univers lointain obtenues dans le cadre du sondage *VIMOS-VLT Deep Survey* (chapitre [5](#)).

1.5.1 Dans l'Univers proche

1.5.1.1 Population globale

La fonction de corrélation à deux points des galaxies proches dans l'espace réel a été déterminée par de nombreux sondages spectroscopiques comme étant une loi de puissance $\xi(r) = (r/r_0)^\gamma$ sur l'intervalle $1 < r < 20 h^{-1} \text{Mpc}$. L'analyse des premières données du

CfA conduisent à une longueur de corrélation $r_0 = 5.4 \pm 0.3 \text{ h}^{-1} \cdot \text{Mpc}$ et une pente $\gamma = 1.77$ (Davis et Peebles, 1983).

L'utilisation combinée des données du CfA2 et du SSRS conduisent à des résultats similaires avec $r_0 = 4.95 \pm 0.13 \text{ h}^{-1} \cdot \text{Mpc}$ et $\gamma = 1.73 \pm 0.03$ lorsque l'on ne tient pas compte des amas (Marzke et al., 1995). De même, Jing et al. (1998) mesurent une longueur de corrélation de $5.06 \pm 0.12 \text{ h}^{-1} \cdot \text{Mpc}$ et une pente $\gamma = 1.862 \pm 0.034$ à partir de 19558 galaxies du LCRS, alors que Tucker et al. (1997) mesurent une longueur de corrélation, dans l'espace des *redshifts* de $s_0 = 6.28 \pm 0.27 \text{ h}^{-1} \cdot \text{Mpc}$ et une pente $\gamma = 1.52 \pm 0.03$.

Lorsque l'on fait des comparaisons, il est important de faire attention à ce que les échelles de distances soient exprimées dans la même cosmologie. Les plus vieilles publications utilisent principalement une cosmologie où $\Lambda = 0$ alors que les recherches les plus récentes placent le contexte dans un Univers plat avec une cosmologie ΛCDM . Or les distances ne s'expriment pas de la même façon d'une cosmologie à l'autre.

À plus grande longueur d'onde, en infrarouge, Fisher et al. (1994) mesurent une valeur de $r_0 = 3.76 \text{ h}^{-1} \cdot \text{Mpc}$ et une pente de 1.66 pour $r < 20 \text{ h}^{-1} \cdot \text{Mpc}$ dans le cadre du sondage IRAS à 1.2 Jy. De même, en utilisant les données du sondage QDOT, Saunders et al. (1992) mesurent des valeurs similaires avec $r_0 = 3.79 \pm 0.14 \text{ h}^{-1} \cdot \text{Mpc}$ et $\gamma = 1.57 \pm 0.03$.

À partir de deux champs disjoints du sondage SSRS2, Willmer et al. (1998) mesurent l'agrégation de galaxies plus brillantes que $m_B = 15.5$. Le premier champ de 3573 galaxies du ciel austral conduit à une longueur de corrélation de $r_0 = 5.12 \pm 0.37 \text{ h}^{-1} \cdot \text{Mpc}$ alors que celle de l'échantillon de l'hémisphère nord, qui contient 1939 galaxies, est de $r_0 = 6.01 \pm 0.44 \text{ h}^{-1} \cdot \text{Mpc}$. Ces deux valeurs mesurées dans deux zones du ciel indépendantes mettent en avant l'effet de la variance cosmique.

Dans le cadre du sondage Durham/UKST, Ratcliffe et al. (1998b) mesurent une valeur de $r_0 = 5.1 \pm 0.3 \text{ h}^{-1} \cdot \text{Mpc}$, identique aux incertitudes près à celle de Loveday et al. (1995) dans le cadre du sondage Stromlo-APM, ($r_0 = 5.1 \pm 0.2 \text{ h}^{-1} \cdot \text{Mpc}$).

On pourra citer aussi l'étude menée par Guzzo et al. (2000) dans le cadre du projet ESP. Pour l'échantillon complet en magnitude à $b_J = 19.4$, la fonction de corrélation $\xi(r)$ est modélisée correctement en loi de puissance avec $r_0 = 4.15^{+0.20}_{-0.21} \text{ h}^{-1} \cdot \text{Mpc}$ et une pente $\gamma = 1.67^{+0.07}_{-0.09}$.

En utilisant le premier jeu de données du SDSS, soit 29300 galaxies couvrant une région du ciel de 690 degrés carrés, Zehavi et al. (2002) mesurent une fonction de corrélation moyenne pour $0.1 < r < 16 \text{ h}^{-1} \cdot \text{Mpc}$ sous la forme

$$\xi(r) = \left(\frac{6.1 \pm 0.2 \text{ h}^{-1} \cdot \text{Mpc}}{r} \right)^{1.75 \pm 0.03} \quad (1.23)$$

Cette expression est très proche de celle de Tucker et al. (1997). En utilisant des données plus complètes du SDSS, avec cette fois 118149 galaxies, Zehavi et al. (2004) mesurent

$$\xi(r) = \left(\frac{5.77 \text{ h}^{-1} \cdot \text{Mpc}}{r} \right)^{1.80} \quad (1.24)$$

Il s'agit d'un résultat remarquable même si le point le plus important de cette étude est la démonstration claire d'un changement de la pente de la fonction de corrélation $\xi(r)$ à $1.5 \text{ h}^{-1} \cdot \text{Mpc}$.

Dans le cadre du 2dFGRS, Hawkins et al. (2003) mesurent une longueur de corrélation de $r_0 = 5.05 \pm 0.26 \text{ h}^{-1} \cdot \text{Mpc}$, plus faible que celle mesurée par dans le SDSS. Cette

différence s'explique, selon [Hawkins et al. \(2003\)](#), par les fonctions de sélection différentes des deux sondages. Les galaxies du SDSS sont sélectionnées en R alors que celles du 2dFGRS le sont en bande B.

1.5.1.2 Les amas

Les amas de galaxies permettent de sonder l'Univers sur de très grandes échelles. [Bahcall et Soneira \(1983\)](#) mesurent l'agrégation des amas d'[Abell \(1958\)](#) et montrent qu'ils sont fortement corrélés jusqu'à des échelles de $\sim 150 h^{-1} \text{ Mpc}$. La forme de la fonction de corrélation des amas est la même que celle des galaxies mais avec une amplitude 18 fois supérieure. La longueur de corrélation mesurée est en effet de l'ordre de $\sim 25 h^{-1} \text{ Mpc}$. Les amas de galaxies se trouvent localisés dans les plus importants pics du champ de densité primordial ce qui explique cette différence.

Le catalogue d'Abell, où les amas ont été identifiés à l'oeil, souffre de nombreux biais. La première réelle alternative à ce catalogue fut celui issu du sondage de galaxies APM. [Dalton et al. \(1994\)](#) mesurent à partir de ce nouveau catalogue une longueur de corrélation $r_0 = 14.3 h^{-1} \text{ Mpc}$.

Une autre méthode d'identification d'amas riches est basée sur leur émissivité en rayons X. Celle-ci est engendrée par le gaz chaud piégé à l'intérieur du puits de potentiel gravitationnel créé par l'amas. Ce rayonnement X est aussi relié à la masse totale de l'amas. Ainsi, dans le cadre du sondage REFLEX, [Borgani et Guzzo \(2001\)](#) mesurent des longueurs de corrélation $\sim 20 h^{-1} \text{ Mpc}$.

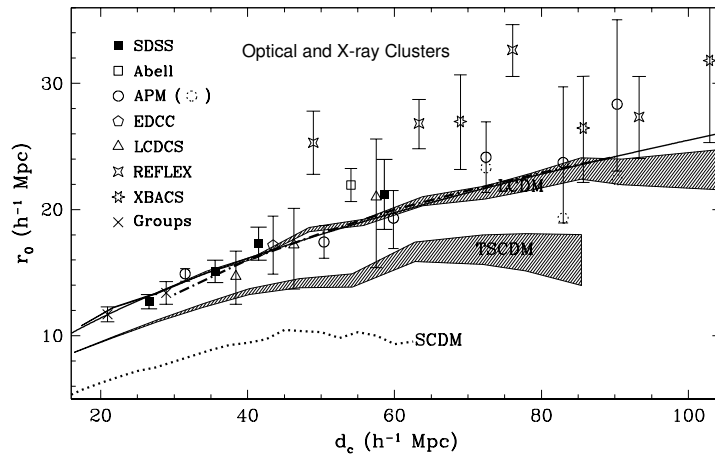


FIG. 1.13 – Longueur de corrélation d'amas de galaxies sélectionnés en optique et en X dans différents sondages représentés en fonction de la séparation moyenne entre amas. Le modèle analytique $r_0 \simeq 2.6\sqrt{d}$ est représenté par la ligne continue. Les zones grisées correspondent à des simulations. Figure 2 de [Bahcall et al. \(2003\)](#)

L'amplitude de l'agrégation des amas dépend de leur richesse, définie en terme de densité, ce que confirment [Bahcall et al. \(2003\)](#) et [Basilakos et Plionis \(2004\)](#). À partir d'amas identifiés dans le SDSS, la longueur de corrélation des amas riches mesurée par [Basilakos et Plionis \(2004\)](#) est de $r_0 = 20.7^{+4.0}_{-3.8} h^{-1} \text{ Mpc}$ contre $r_0 = 9.7^{+1.2}_{-1.2} h^{-1} \text{ Mpc}$.

[Bahcall et al. \(2003\)](#) concluent que cette dépendance est en accord avec le modèle cosmologique standard Λ CDM ($\Omega_m=0.3$, $h=0.7$) et par contre complètement inconsistent avec des modèles où $\Omega_m=1$ qui prédisent des longueurs de corrélations plus faibles. Ils dérivent enfin une relation analytique entre la longueur de corrélation r_0 et la séparation moyenne des amas : $r_0 \simeq 2.6\sqrt{d}$. La figure 1.13 décrit ce modèle comparé à diverses observations et simulations.

1.5.1.3 En fonction de la luminosité

Dès la fin des années 80 il est fait état d'une ségrégation des galaxies en fonction de leur luminosité. En effet, [Hamilton \(1988\)](#) observe dans les données du CfA que les galaxies brillantes sont 3.5 fois plus agrégées que les galaxies faibles.

Dans le cadre du sondage Stromlo-APM, [Loveday et al. \(1995\)](#) divisent leur échantillon de 1757 galaxies en fonction de leurs magnitudes absolues en B. Ils définissent ainsi trois échantillons d'objets dont la magnitude appartient à $[-22 ; -20]$, $[-20 ; -19]$ ou $[-19 ; -15]$. Les galaxies faibles sont ainsi ~ 2 fois moins agrégées que les galaxies de type L^* et brillantes (valeurs rassemblées dans la table 1.1).

[Guzzo et al. \(1997\)](#) créent 3 catalogues de magnitude absolue limite croissante à partir des données du sondage *Perseus-Pisces survey*. Ils détectent une augmentation de 25% de la longueur de corrélation r_0 entre l'échantillon limité à $M_{Zw} = -19$ et celui limité à -20 . En utilisant la même technique, à savoir en définissant des échantillons limités en volume, [Benoist et al. \(1996\)](#) et [Willmer et al. \(1998\)](#) observent cette même ségrégation des galaxies en fonction de leur luminosité dans les données complétées du SSRS2. En s'affranchissant de la distorsion de la distribution des galaxies dans l'espace des *redshifts* (voir partie 3.1.1), [Guzzo et al. \(2000\)](#) mesurent une augmentation de la longueur de corrélation de $\sim 50\%$ lorsque la limite en magnitude absolue passe de -18.5 à -20 dans le cadre du sondage *ESO Slice Project*.

Plus récemment, dans le cadre du CNOC2, l'un des premiers sondages spectroscopiques à atteindre des *redshifts* de l'ordre de $z \sim 0.5$, [Shepherd et al. \(2001\)](#) rapportent une dépendance de l'agrégation des galaxies en fonction de la luminosité des galaxies. En ne considérant que les galaxies dont le *redshift* est compris entre 0.12 et 0.4, ils divisent leur échantillon de galaxies en quatre intervalles de magnitudes absolues. En normalisant leurs résultats à une loi de puissance $\xi(r)$ de pente 1.73, ils mesurent une augmentation de r_0 de 3.32 à 4.25 h^{-1} .Mpc avec la luminosité.

Afin de clarifier à $z \sim 0$ la dépendance de l'agrégation des galaxies en fonction de la luminosité, des études ont été menées à partir des données des sondages 2dFGRS et SDSS. La qualité et la quantité de données de ces deux projets sont telles qu'elles permettent de considérer un plus grand nombre de paramètres et de mesurer l'état d'agrégation des galaxies de façon beaucoup plus précise. Ainsi, en calculant la fonction de corrélation angulaire, [Budavári et al. \(2003\)](#) mesurent une longueur de corrélation de 5.04 h^{-1} .Mpc pour des galaxies plus faibles que $M_{r^*} = -21$, alors que celle de leur échantillon le plus brillant est de 7.87 h^{-1} .Mpc. Peu avant, en utilisant les premières données du SDSS ($5700 \text{ km.s}^{-1} \leq cz \leq 39000 \text{ km.s}^{-1}$, 29300 galaxies), [Zehavi et al. \(2002\)](#) mettaient en évidence de façon très nette la dépendance à la luminosité en calculant la projection de la fonction de corrélation spatiale $\xi(r_p, \pi)$ (partie 3.1.1). Ce résultat est indiqué par la figure 1.14 où l'amplitude de la fonction $w_p(r_p)$ augmente avec la luminosité.

Dans une analyse plus récente des données du SDSS avec ~ 200000 galaxies, [Zehavi et al. \(2005b\)](#) confirment ce résultat (figure 1.15).

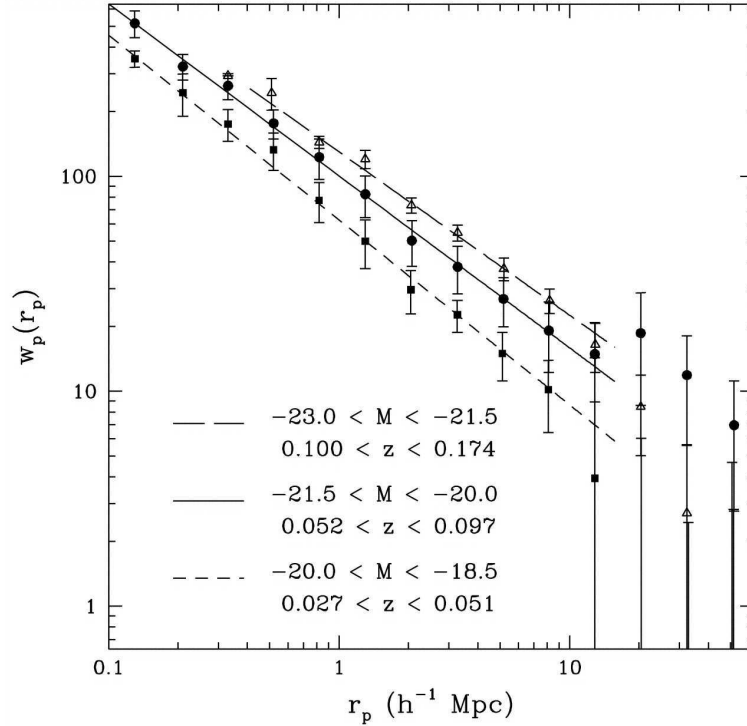


FIG. 1.14 – $w_p(r_p)$ en fonction de la luminosité calculée dans le SDSS ([Zehavi et al., 2002](#), figure 16)

Enfin, en utilisant 160000 *redshifts* du 2dFGRS, [Norberg et al. \(2001\)](#) observent la même tendance, avec des galaxies brillantes plus fortement agrégées que les galaxies faibles.

Grâce au développement des simulations numériques, la dépendance de l'agrégation des galaxies à la luminosité est maintenant attendue dans un Univers hiérarchique *Cold Dark Matter* ([Benson et al., 2001](#)) ainsi qu'illustré par la figure 1.16. En effet, les galaxies les plus lumineuses se forment de préférence dans les halos de matière noire les plus massifs. Or ceux-ci sont fortement agrégés.

TAB. 1.1 – Longueur de corrélation r_0 (exprimée en $h^{-1} \text{Mpc}$) en fonction de la luminosité. On ne donne ici que les cas extrêmes de chaque étude

	galaxies faibles		galaxies brillantes	
	sélection	r_0	sélection	r_0
<i>a</i>	$-19 \leq M_B \leq -15$	3.2 ± 0.6	$-22 \leq M_B \leq -20$	5.8 ± 0.3
<i>b</i>	$M_{Zw} < -19$	$5.95^{+0.27}_{-0.31}$	$M_{Zw} < -20$	$7.05^{+0.47}_{-0.59}$
<i>c</i>	$M_B < -17$	4.3 ± 0.5	$M_B < -21$	15.8 ± 2.9
<i>d</i>	$M_B < -18.39$	4.17 ± 0.24	$M_B < -19.89$	7.47 ± 0.42
<i>e</i>	$M_B < -18.5$	$3.45^{+0.21}_{-0.30}$	$M_B < -20$	$5.15^{+0.39}_{-0.44}$
<i>f</i>	$-19.65 < M_R < -19.25$	3.32 ± 0.44	$-22.52 < M_R < -20.60$	4.25 ± 0.28
<i>g</i>	$M_{r^*} > -21$	5.04 ± 0.13	$-23 < M_{r^*} < -22$	7.87 ± 0.33
<i>h</i>	$-20 < M_{r^*} < -18.5$	4.72 ± 0.44	$-23 < M_{r^*} < -21.5$	7.42 ± 0.33
<i>i</i>	$-18.5 < M_{bj} < -18$	4.14 ± 0.64	$-21.5 < M_{bj} < -21$	7.58 ± 0.48

a Stromlo-APM ; Loveday *et al.* (1995) ; $\Lambda=0$, $q_0 = 0.5$ et $h=1$

b Perseus-Pisces ; Guzzo *et al.* (1997) ; $h=1$

c SSRS2 ; Benoist *et al.* (1996) ; $q_0 = 0.5$ et $h=1$

d SSRS2 ; Willmer *et al.* (1998) ; $h=1$

e ESP ; Guzzo *et al.* (2000) ; $q_0 = 0.5$ et $h=1$

f CNOC2 ; Shepherd *et al.* (2001) ; $0.12 < z < 0.4$, $\Omega_m=0.2$, $\Omega_\Lambda=0$, $h=1$

g SDSS ; Budavári *et al.* (2003) ; $r^* < 21$, $\Omega_m=0.3$, $\Omega_\Lambda=0.7$, $h=0.7$

h SDSS ; Zehavi *et al.* (2002) ; $\Omega_m=0.3$, $\Omega_\Lambda=0.7$, $h=0.7$

i 2dFGRS ; Norberg *et al.* (2001) ; $\Omega_m=0.3$, $\Omega_\Lambda=0.7$, $h=1$

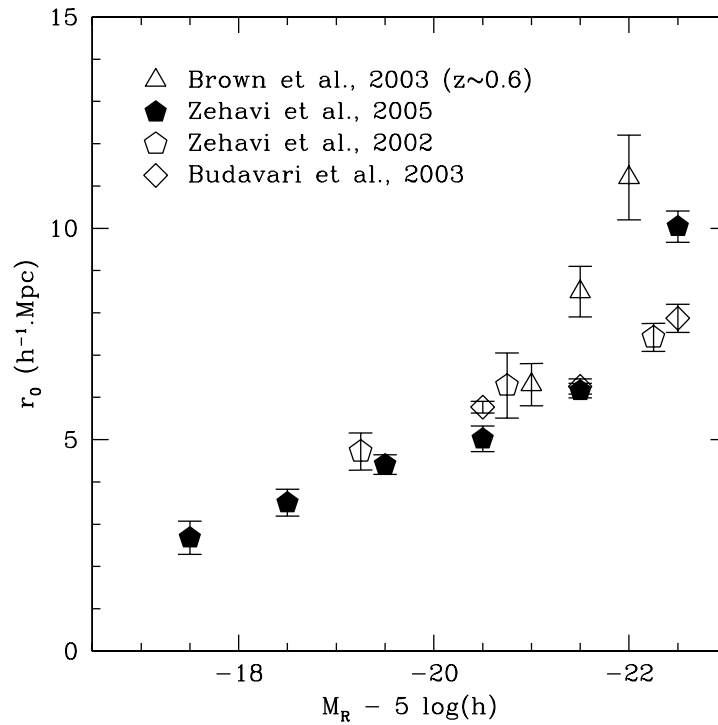


FIG. 1.15 – Longueur de corrélation dans l'espace réel en fonction de la magnitude absolue en bande R

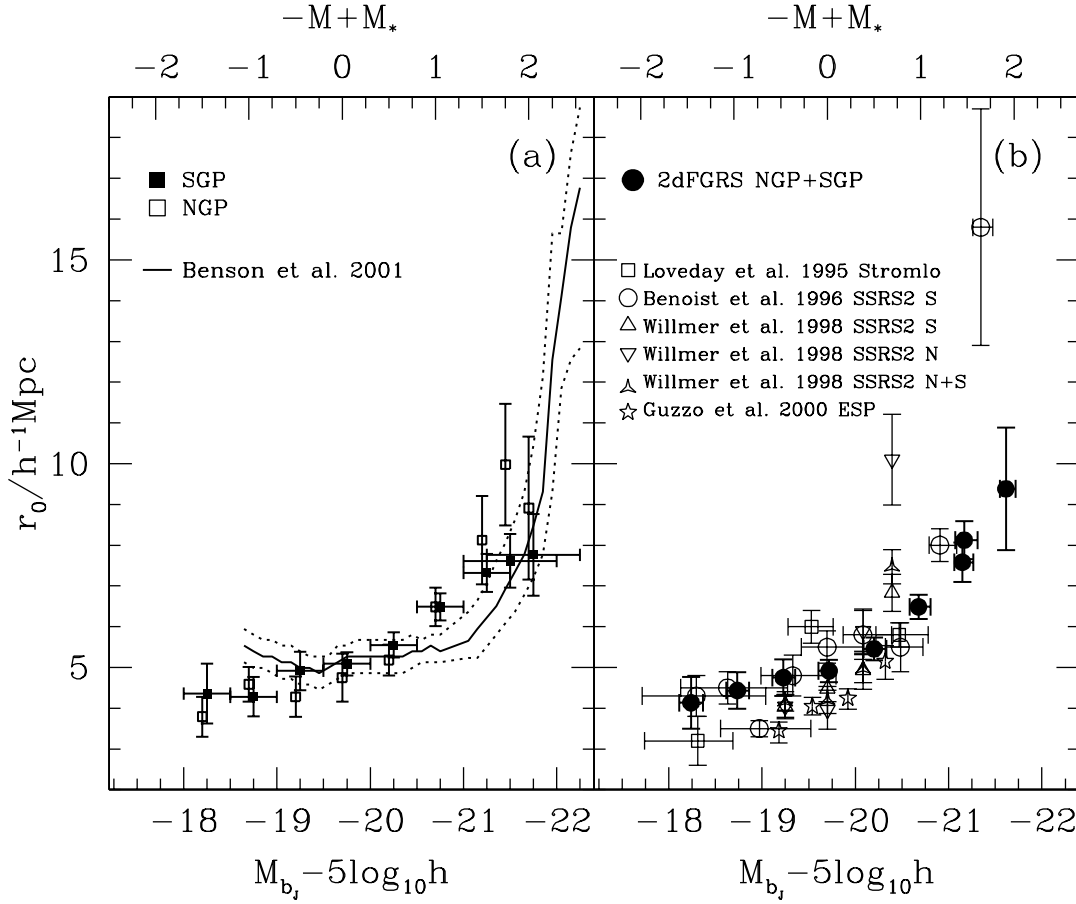


FIG. 1.16 – Longueur de corrélation en fonction de la magnitude absolue en bande B (Norberg *et al.*, 2001, figure 3). Le panneau (a) montre les prédictions du modèle semi-analytique de Benson *et al.* (2001) (trait plein) avec une estimation des erreurs associées (traits pointillés) confronté aux mesures obtenues dans le 2dF nord et sud. Le panneau (b) indique les résultats combinés du 2dFGRS avec d'autres résultats récents.

1.5.1.4 En fonction du type morphologique

Hubble (1936) remarqua que les galaxies ont des distributions spatiales différentes en fonction de leur morphologie. Les galaxies précoces (*early-type*), telles que les galaxies elliptiques ou lenticulaires, sont fortement concentrées et comptent pour une grande partie des galaxies au centre des amas riches, alors que les galaxies tardives (*late-type*), de type spirales ou irrégulières, sont plus diffusées et représentent jusqu'à 70% des galaxies de champ (Dressler, 1980).

La première tentative de quantifier la différence entre l'agrégation des galaxies de types précoces et celles de types tardifs remonte à l'analyse du catalogue Uppsala par Davis et Geller (1976). L'amplitude de la fonction de corrélation angulaire $\omega(\theta)$ des galaxies elliptiques était dès lors plus importante que celle des galaxies de type spiral, avec une pente plus grande. Giovanelli *et al.* (1986) observent ce même effet dans les données du sondage *Perseus-Pisces survey*, avec cette même relation densité-morphologie

(Postman et Geller, 1984).

En étudiant la vitesse circulaire des galaxies du *Nearby Galaxy Catalogue* dans les amas, White *et al.* (1988) observent, comme prévu théoriquement dans le cadre d'un modèle *Cold Dark Matter* que l'agrégation de ces galaxies est une fonction croissante de cette vitesse particulière. Les galaxies présentant une grande vitesse se trouvent préférentiellement dans les régions très denses. Ils relient cet effet à la relation morphologie-densité de Dressler (1980).

Plusieurs études plus récentes, réalisées indépendamment à partir de différents sondages, confirment ces premières investigations. Santiago et Strauss (1992) constatent par exemple dans le CfA que les galaxies elliptiques et spirales sont deux fois plus agrégées que les irrégulières et que les elliptiques sont d'avantage concentrées que les galaxies spirales dans les amas *Virgo* et *Ursa Major*. Iovino *et al.* (1993) observent cette même ségrégation en fonction du type des galaxies dans le champ *Perseus-Pisces*. En mesurant la fonction de corrélation à deux points dans l'espace des *redshifts*, $\xi(s)$, à partir des données du sondage ORS, Hermit *et al.* (1996) mesurent des longueurs de corrélation $s_0 = 5.65 \text{ h}^{-1} \cdot \text{Mpc}$ pour les galaxies précoces, $4.24 \text{ h}^{-1} \cdot \text{Mpc}$ pour les galaxies spirales S_{ab} , et $3.78 \text{ h}^{-1} \cdot \text{Mpc}$ pour les galaxies tardives. Dans le cadre du sondage *Stromlo-APM*, Loveday *et al.* (1995) observent eux aussi cette ségrégation en divisant leur échantillon de galaxies en deux (elliptiques/lenticulaires et spirales/irrégulières) avec des magnitudes absolues en bande B comprises entre -22 et -15 .

Guzzo *et al.* (1997) font remarquer que les galaxies elliptiques/lenticulaires étant visiblement plus concentrées au centre des amas, leur vitesse particulière est en moyenne plus importante que celle des galaxies de champs. Il en découle que dans l'espace des *redshifts*, la fonction de corrélation de cette classe de galaxie est très affectée à petite échelle par une perte du signal mesurable. La solution est donc soit de calculer la fonction de corrélation angulaire, soit la fonction de corrélation $\xi(r_p, \pi)$ que l'on intégrera selon la ligne de visée pour s'affranchir des effets de distorsion (partie 3.1.1). Ainsi, ces auteurs observent, là encore dans le champ *Perseus-Pisces*, que même pour des galaxies de magnitudes plus brillantes que -19.5 en bande B, les galaxies elliptiques sont plus agrégées que les spirales (valeurs rassemblées table 1.2). C'est aussi le cas des résultats obtenus par Willmer *et al.* (1998) à partir des dernières données du SSRS, mais ici l'échantillon considéré va jusqu'à une magnitude de -13 . En considérant un échantillon limité en volume comme celui de Guzzo *et al.* (1997), les longueurs de corrélations mesurées dans le SSRS restent inférieures à celles du *Perseus-Pisces*, interprétées par Willmer *et al.* (1998) comme un manque d'amas riches dans leurs données et donc l'effet de la variance cosmique. Ces dernières permettent tout de même de constater que l'état d'agrégation des deux types de galaxies du SSRS augmentent avec la luminosité.

1.5.1.5 En fonction de la couleur/type spectral

Ségrégation en couleur

Il a très rapidement été observé une ségrégation des galaxies selon qu'elles soient plutôt bleues ou rouges. Quelques exemples de sélections de galaxies bleues et rouges sont décrites dans la partie 1.4.2.3.

Neuschaefer *et al.* (1995) observent ainsi que l'amplitude de $\omega(\theta)$ de la population

TAB. 1.2 – Longueur de corrélation r_0 (exprimée en $h^{-1} \text{Mpc}$) en fonction du type morphologique des galaxies. On ne donne ici que les cas extrêmes de chaque étude

sondage	r_0	
	galaxies tardives	galaxies précoces
Stromlo-APM ^a	4.4 ± 0.1	5.9 ± 0.7
Perseus-Pisces ^b	$5.55^{+0.40}_{-0.45}$	$8.35^{+0.75}_{-0.76}$
SSRS2 ^c	5.26 ± 0.37	6.06 ± 0.39
ORSm ^d	4.71	6.70
ORSd ^d	3.78	5.65
SDSS ^e	5.64 ± 0.22	6.74 ± 0.24

^aLoveday *et al.* (1995) ; $\Lambda=0$, $q_0 = 0.5$ et $h=1$

^bGuzzo *et al.* (1997) ; magnitudes absolues calculées avec $h=1$

^cWillmer *et al.* (1998) ; magnitudes absolues calculées avec $h=1$

^dHermit *et al.* (1996) ; Les valeurs reportées ici sont celles de s_0 , longueur de corrélation dans l'espace des *redshifts* $\xi(s)$

^eZehavi *et al.* (2002) ; $\Omega_m=0.3$, $\Omega_\Lambda=0.7$, $h=0.7$; classification en terme de concentration

rouge est plus forte que celle de la population bleue. Landy *et al.* (1996) mesurent le même effet. Le résultat est le même d'un point de vue qualitatif quelque soit leur définition "rouge/bleue" utilisée, bien qu'il soit plus prononcé dans le cas d'une couleur (U-R) qui décrit une plus grande plage en longueur d'onde.

La forte amplitude mesurée pour la population rouge est expliquée par la nature de ces galaxies, essentiellement de type elliptique/lenticulaire (relation morphologie/densité, voir section 1.5.1.4). Par contre beaucoup d'interrogations sont soulevées quant à la nature de la population bleue. S'agit-il de précurseurs de galaxies normales observées à grand *redshift* en train de former des étoiles ? S'agit-il de galaxies à moyen/grand décalage spectral en train de fusionner ? S'agit-il d'une population faible à bas *redshift* ?

Roche *et al.* (1996) observent la même différence d'amplitude de la fonction de corrélation angulaire entre celle des galaxies bleues et celle des galaxies rouges. Ces amplitudes sont confrontées à différents modèles, desquels sont exclus ceux ne présentant pas d'évolution.

La faible statistique des données de Willmer *et al.* (1998) induit une mesure bruitée de la longueur de corrélation spatiale r_0 mais celle de la population rouge est nettement supérieure à celle de la population bleue à 2σ (table 1.3).

En utilisant une définition en couleur légèrement plus raffinée (équations 1.14 et 1.15), les propriétés d'agrégation des galaxies de Brown *et al.* (2000) dépendent là encore de la couleur de celles-ci. Les galaxies rouges, associées à des types précoces, sont fortement amalgamées, avec une fonction de corrélation plus pentue que celle des galaxies bleues. Brown *et al.* (2000) concluent que la nature de l'agrégation dépend plus de la couleur et de la population stellaire dominante que de la morphologie.

Dans le cadre du SDSS, Zehavi *et al.* (2002) calculent la fonction de corrélation $\xi(r_p, \pi)$ de leur deux populations et celle-ci montre une très forte distorsion dans l'espace des *redshifts* de la population rouge, avec un *Finger-of-God*³ très prononcé. Pour s'affranchir de

³littéralement le "doigt de Dieu", voir section 3.1.1 pour tout ce qui concerne la fonction de corrélation à deux dimensions $\xi(r_p, \pi)$

ces effets, la fonction $w_p(r_p)$ est calculée (Zehavi *et al.*, 2002, figure 16) et sa modélisation par une loi de puissance fournit les longueurs de corrélation reportées table 1.3. Un changement de la pente de $w_p(r_p)$ pour les galaxies bleues à faible séparation traduirait la tendance pour ces galaxies à se retrouver dans des halos de faibles masses.

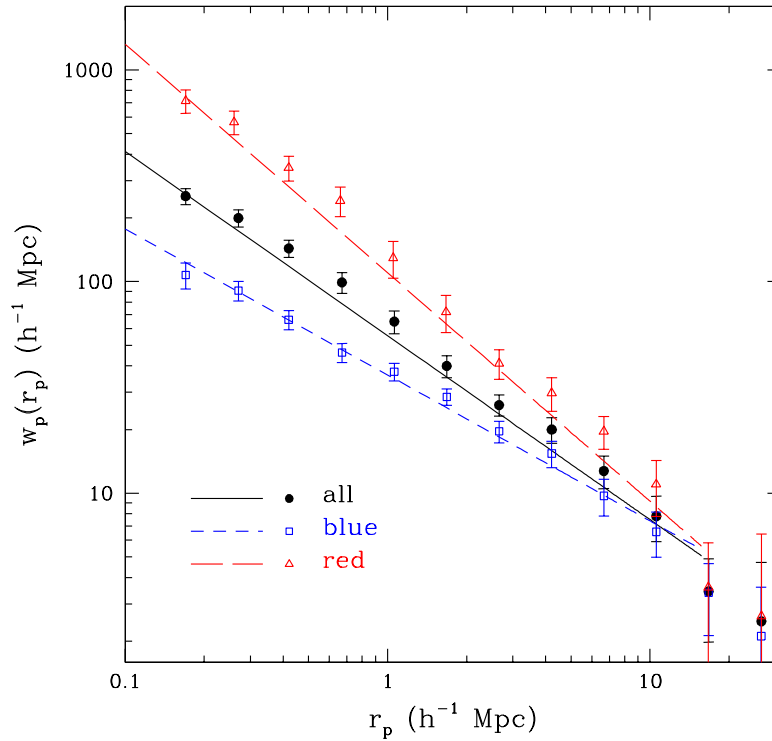


FIG. 1.17 – $w_p(r_p)$ en fonction de la couleur dans le SDSS (Zehavi *et al.*, 2005b, figure 13) dans l'intervalle en magnitude $-20 < M_r < -19$

En utilisant la deuxième époque de données du SDSS (*Data Release 2*), Zehavi *et al.* (2005b) mesurent dans l'intervalle en magnitude $-20 < M_r < -19$, les paramètres de la fonction de corrélation et obtiennent $r_0 = 5.7 \text{ h}^{-1} \text{ Mpc}$ et $\gamma = 2.1$ pour la population rouge et $r_0 = 3.6 \text{ h}^{-1} \text{ Mpc}$ et $\gamma = 1.7$ pour la population bleue (figure 1.17). Les valeurs pour d'autres intervalles de magnitude absolue sont qualitativement similaires : augmentation de la longueur de corrélation et de la pente entre la classe bleue et la rouge.

La figure 1.18 donne une vue synthétique de la ségrégation en couleur des galaxies. L'axe des abscisses est défini arbitrairement du bleu au rouge donc deux points à la même abscisse ne sont donc pas directement comparables. Par contre globalement, les longueurs de corrélation des galaxies rouges sont plus importantes que celles des bleues pour un même jeu de mesures. On retiendra aussi que la couleur étant liée au contenu stellaire/gazeux des galaxies, la fonction de corrélation des galaxies rouges mesure la distribution des galaxies dominées par une population d'étoiles vieilles, alors que celle des galaxies bleues mesure la distribution des galaxies dont le taux de formation stellaire est important.

TAB. 1.3 – Longueur de corrélation r_0 (exprimée en $h^{-1}.\text{Mpc}$) en fonction de la couleur des galaxies. On ne donne ici que les cas extrêmes de chaque étude

sondage	r_0	
	galaxies bleues	galaxies rouges
SSRS2 ^a	6.39 ± 0.50	7.99 ± 0.95
SDSS ^b	4.02 ± 0.25	6.78 ± 0.23
SDSS ^c	3.6	5.7

^aWillmer *et al.* (1998) ; magnitude absolue calculées avec $h=1$

^bZehavi *et al.* (2002) ; $\Omega_m=0.3$, $\Omega_\Lambda=0.7$, $h=0.7$

^cZehavi *et al.* (2005b)

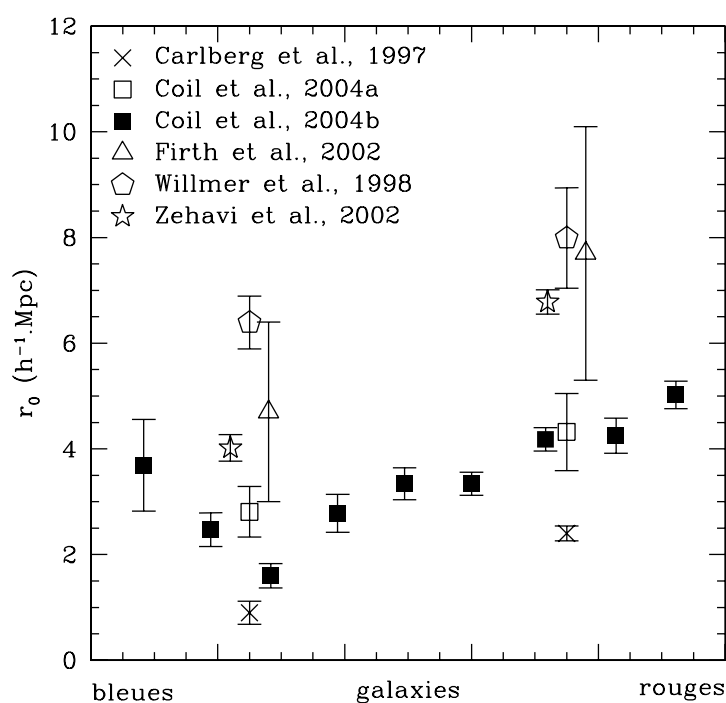


FIG. 1.18 – Compilation de résultats établissant la ségrégation des galaxies en fonction de leur couleur. L'axe des abscisses est défini arbitrairement donc deux points à la même abscisse n'ont pas forcément une couleur comparable. Par contre la tendance d'ensemble indique bien que les galaxies rouges sont les plus agrégées.

Ségrégation en fonction du type spectral

Des classifications qui utilisent non seulement l'information multi-longueur d'onde mais aussi l'information spectrale (partie 1.4.2.3) conduisent à des résultats analogues.

Loveday *et al.* (1999) observent dans les données du sondage *Stromlo-APM* que les objets présentant des raies en émission sont moins agrégés que ceux dont le spectre en est dépourvu. La longueur de corrélation des objets à fortes raies d'émission est ainsi dans le cas extrême plus de deux fois plus faible que celle des objets faisant état de peu d'activité stellaire.

En mesurant les paramètres de la fonction de corrélation $\xi(r)$ dans différents intervalles de magnitudes absolues à partir des données du 2dFGRS, Norberg *et al.* (2002) confirment de façon très précise la ségrégation des galaxies selon leur type spectral. De plus, pour un type de galaxie donné, une ségrégation en luminosité est aussi visible (figure 1.19). En considérant cette fois l'échantillon complet limité en magnitude à $b_J = 19.45$, Madgwick *et al.* (2003) analysent l'agrégation des deux mêmes classes de galaxies, précoces et tardives. La longueur de corrélation mesurée pour la population de galaxies passive est ~ 2 fois plus grande que celle de la population de galaxies actives (table 1.4).

Dans le cadre du SDSS, Zehavi *et al.* (2002) établissent que la fonction de corrélation des galaxies de forte concentration est de plus grande amplitude et plus pentue que celle des galaxies peu concentrées. Les valeurs obtenues par Budavári *et al.* (2003) en calculant la fonction de corrélation angulaire confirment là encore la ségrégation des galaxies en fonction de leur type, mais aussi de leur luminosité.

La figure 1.20 montre différentes valeurs de longueurs de corrélation r_0 mesurées par différents auteurs en fonction du type de chaque galaxie. La définition des types donnée sur cette figure est arbitraire. Deux valeurs à la même abscisse ne correspondent pas forcément à un type identique de galaxies. Par contre le résultat d'ensemble est bien marqué, avec des galaxies précoces plus fortement agrégées que les galaxies de type plus tardif.

TAB. 1.4 – Longueur de corrélation r_0 (exprimée en $h^{-1} \text{Mpc}$) en fonction du type spectral des galaxies. On ne donne ici que les cas extrêmes de chaque étude

sondage	r_0	
	galaxies tardives	galaxies précoces
Stromlo-APM ^a	2.9 ± 0.7	6.2 ± 1.8
2dFGRS ^b	4.27 ± 0.81	8.33 ± 1.82
2dFGRS ^c	6.33 ± 1.01	7.60 ± 1.02
2dFGRS ^d	3.67 ± 0.30	6.10 ± 0.34
SDSS ^e	4.51 ± 0.25	6.59 ± 0.24
SDSS ^f	5.52 ± 0.68	7.91 ± 0.48

^aLoveday *et al.* (1999) ; $q_0 = 0.5$ et $h=1$

^bNorberg *et al.* (2002) ; $\Omega_m=0.3$, $\Omega_\Lambda=0.7$, $h=1$; $-18.5 < M_{b_J} - 5 \log(h) < -17.5$

^cNorberg *et al.* (2002) ; $\Omega_m=0.3$, $\Omega_\Lambda=0.7$, $h=1$; $-21.5 < M_{b_J} - 5 \log(h) < -20.5$

^dMadgwick *et al.* (2003) ; échantillon complet à $b_J \leq 19.45$ et $0.01 < z < 0.15$

^eBudavári *et al.* (2003) ; $r^* < 21$, $\Omega_m=0.3$, $\Omega_\Lambda=0.7$, $h=0.7$; $M_{r^*} > -21$

^fBudavári *et al.* (2003) ; $r^* < 21$, $\Omega_m=0.3$, $\Omega_\Lambda=0.7$, $h=0.7$; $-23 < M_{r^*} < -21$

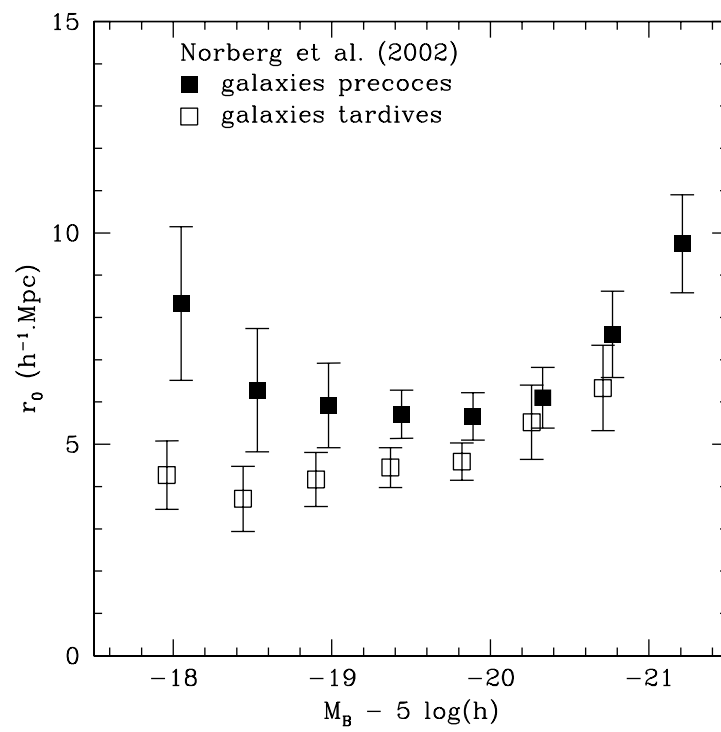


FIG. 1.19 – Longueur de corrélation r_0 en fonction de la magnitude absolue en bande b_J pour chacun des deux types de galaxies définis par [Norberg et al. \(2002\)](#) (adaptation de la figure 8 de cet article)

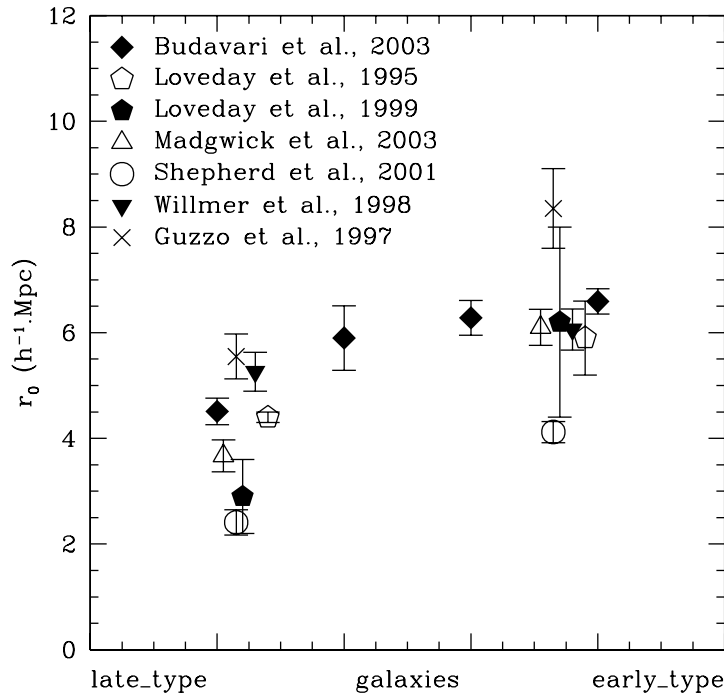


FIG. 1.20 – Longueur de corrélation r_0 en fonction du type de chaque galaxie. La définition de l'axe des abscisses est arbitraire. Deux valeurs à la même abscisse ne correspondent pas forcément à un type identique de galaxies. Par contre le résultat d'ensemble est bien marqué, avec des galaxies précoces plus fortement agrégées que les galaxies de type plus tardif.

1.5.1.6 Synthèse

Beaucoup d'études sur l'état d'agrégation des galaxies ont été menées dans l'Univers proche. La mesure de la fonction de corrélation spatiale et/ou angulaire montre des ségrégations dans la distribution des galaxies. Ainsi, on retiendra que les galaxies rouges sont plus agrégées que les bleues, les lumineuses plus que les faibles, les types précoces plus que les types tardifs. La pente de la fonction de corrélation spatiale à deux points est généralement plus importante dans le cas des galaxies rouges, lumineuses et/ou de types précoces que leurs contreparties bleues, faibles, et de types tardifs.

Les galaxies précoces de type elliptique ont une population d'étoiles plus vieilles et plus rouges que les galaxies tardives de type spirale ou irrégulière bleue, à fort taux de formation d'étoiles. Ces dernières sont essentiellement des galaxies de champs alors que les elliptiques se trouvent au centre des amas, et associées à des halos massifs de matière noire.

Les modèles hiérarchiques de formation des galaxies prédisent ces ségrégations et coïncident très bien avec les observations récentes. La distribution des galaxies massives montre de fort biais par rapport à celle de la distribution de masse sous-jacente.

1.5.2 À grand décalage spectral

1.5.2.1 Population globale

Afin de suivre l'évolution de l'agrégation des galaxies jusqu'à des *redshifts* de l'ordre de $z \sim 1$, plusieurs projets ont été réalisés. Ils ne sont pas toujours directement comparables entre eux pour des raisons de fonctions de sélection différentes mais ils apportent tout de même leurs lots d'informations quant à l'évolution des grandes structures.

L'un de ces premiers sondages fut celui du *Canada - France Redshift Survey*. À partir de 591 *redshifts* de galaxies jusqu'à $z \sim 1.3$ réparties sur 5 champs de $10' \times 10'$, [Le Fèvre et al. \(1996\)](#) suivent l'évolution des paramètres de la fonction de corrélation spatiale $\xi(r)$. Leur échantillon, qui est sélectionné à $I_{AB} = 22.5$, montre une diminution d'un facteur ~ 10 de l'amplitude de $\xi(r)$ à $z \sim 0.6$ par rapport aux observations faites dans l'Univers local pour des galaxies similaires. Ces auteurs n'observent pas de ségrégation en couleur rouge/bleue à $z \geq 0.5$ mais une différence est détectée pour des *redshifts* intermédiaires ($0.2 \leq z \leq 0.5$) comme dans l'Univers local.

En utilisant des données plus profondes ($I_{AB} = 24$) sur un champ plus grand de 4×4 degrés carrés [Postman et al. \(1998\)](#) mesurent la fonction de corrélation angulaire $\omega(\theta)$. En utilisant 4 modèles de distribution en *redshifts* $N(z)$ dont un basé sur les données du CFRS, ces auteurs mesurent une plus grande longueur de corrélation que [Le Fèvre et al. \(1996\)](#) avec $r_0(z=0.5) \simeq 3.0 \pm 0.4 \text{ h}^{-1} \text{ Mpc}$. Cette différence se justifierait par la taille angulaire des champs, celui du CFRS étant plus sujet à des effets de variance cosmique (les deux sondages sont sélectionnés en I_{AB}). De tels sondages permettent aussi de détecter des structures, amas, à grands *redshifts*. Ainsi [Le Fèvre et al. \(1996\)](#) détectent un amas à $z=0.78$ et une structure dense à $z=0.985$ ([Le Fèvre et al., 1994b](#)).

La nature des galaxies sondées évoluant avec le *redshift* dans le cas d'un sondage limité en flux (les objets à grand décalage spectral sont plus brillants que ceux à bas *redshift*), [Carlberg et al. \(2000\)](#) suivent l'évolution des propriétés d'agrégation de galaxies plus brillantes que $M_R \leq -20$, définissant ainsi un échantillon limité en volume entre $z=0.1$ et $z=0.65$ à partir des données du CNOC2 (voir aussi [Shepherd et al. \(2001\)](#)). Les valeurs de r_0 ainsi mesurées diminuent légèrement avec le *redshift* tout en restant autour de $4 \text{ h}^{-1} \text{ Mpc}$. Dans cette cosmologie ($\Omega_m=0.2$, $\Omega_\Lambda=0.8$), ces mesures sont consistantes avec des modèles d'agrégation biaisées des galaxies alors qu'une cosmologie avec $\Omega_m=1$ semble exclue.

À cette même période, une mesure de l'évolution de l'agrégation des galaxies jusqu'à $z \sim 3.5$ a été menée grâce aux données très profondes des champs HDF-Nord et HDF-Sud ([Arnouts et al., 1999, 2002](#)). La profondeur des données ($I_{AB} = 27.5$) permet en effet de sélectionner des objets très faibles à très grand décalage spectral. La technique des *redshifts* photométriques, confrontés à quelques *redshifts* spectroscopiques, permet de calculer la fonction de corrélation angulaire $\omega(\theta)$ dans 6 intervalles en *redshifts* consécutifs. L'amplitude du biais augmente avec le *redshift* jusqu'à $b \simeq 5$ à $z \simeq 4$. Cette observation supporte un modèle de formation biaisée des galaxies où les galaxies à grands *redshifts* se forment préférentiellement dans les halos massifs. Par contre des études similaires sur de plus grands champs sont nécessaires pour s'affranchir de la variance cosmique.

Si l'on revient à l'intervalle $0.3 < z < 0.9$, [Brown et al. \(2003\)](#) mesurent de fortes longueurs de corrélation pour les galaxies rouges dans le cadre du sondage *NOAO Deep Wide-Field Survey* (NDWFS). Le champ mesure 1.2 degrés carrés et les galaxies étudiées ont

une couleur VEGA au repos de $B_W - R > 1.44$ et une magnitude absolue $M_R - 5 \log(h)$ comprise entre -21.5 et -20.5 . Ces objets rouges sont très agrégés ($r_0 \sim 6$) mais sans évolution très prononcée avec le *redshift*. Ces valeurs importantes vont dans le sens d'une ségrégation en couleur observée dans l'Univers proche. Une ségrégation en luminosité est aussi visible.

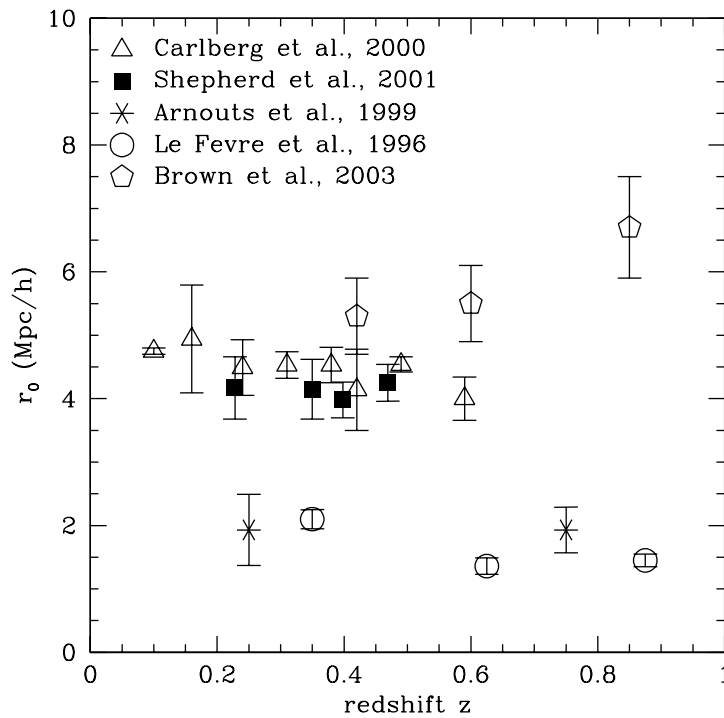


FIG. 1.21 – Longueur de corrélation r_0 en fonction du décalage spectral z . La large dispersion des différentes mesures se justifie par des fonctions de sélection différentes (voir texte).

La figure 1.21 donne une vue synthétique de ces différentes mesures réalisées à grand *redshift*. Les mesures de Le Fevre et al. (1996) et Arnouts et al. (1999) sont plus faibles que les autres mais s'expliquent par des champs de vue petits fortement dépendants de la variance cosmique. De plus, les objets du HDF sont très faiblement lumineux.

1.5.2.2 Les galaxies à discontinuité Lyman

La statistique de ces objets LBGs (définis partie 1.4.2.4) augmente progressivement. Steidel et al. (1998) confirment 78 galaxies dans l'intervalle en *redshifts* $[2.0 - 3.4]$ dont 59 dans $[2.7 - 3.4]$. L'étude de ces objets implique qu'ils soient des traceurs fortement biaisés de la distribution de la masse à ces *redshifts* car ils observent une forte surdensité qui n'existerait pas dans un modèle *Cold Dark Matter* sans cette hypothèse. Ces LBGs sont associées à des halos de matière noire de masse $10^{12} M_\odot$ et sont déjà considérés comme des progéniteurs de galaxies massives observées dans l'Univers local.

À partir de comptages en cellules de 268 LBGs confirmées spectroscopiquement dans

6 champs de $9' \times 9'$ à $z \sim 3$, [Adelberger et al. \(1998\)](#) mesurent un biais compris entre ~ 2 et ~ 6 selon la cosmologie utilisée ($b = 4.0 \pm 0.7$ pour un Univers plat avec $\Omega_m = 0.3$) à des échelles de $\sim 10 \text{ h}^{-1} \text{ Mpc}$. À partir de la variance du nombre de galaxies présentes dans des cellules de ces dimensions, ils mesurent une longueur de corrélation spatiale comprise entre 4 et $6 \text{ h}^{-1} \text{ Mpc}$ selon la cosmologie, avec l'hypothèse d'une pente $\gamma = 1.8$.

[Giavalisco et al. \(1998\)](#) mesurent directement la fonction de corrélation angulaire $\omega(\theta)$ de 871 LBGs réparties dans 5 champs. La distribution en *redshifts* établie à partir de 376 mesures spectroscopiques permet de remonter à la longueur de corrélation spatiale via les équations de [Limber \(1953\)](#). Ils mesurent $r_0 = 3.3^{+0.7}_{-0.6} \text{ h}^{-1} \text{ Mpc}$ pour un facteur de décélération $q_0 = 0.1$, et $r_0 = 2.1^{+0.4}_{-0.5} \text{ h}^{-1} \text{ Mpc}$ pour $q_0 = 0.5$, à un *redshift* médian de 3.04. Ces valeurs sont comparables à celles de galaxies spirales à $z \sim 0$, et légèrement inférieures à celles de galaxies elliptiques. Le biais linéaire mesuré est de 1.5 (4.5) pour $q_0 = 0.1$ (0.5). Ces mesures sont en accord avec la théorie de formation biaisée des galaxies. La technique LBG sélectionne des galaxies brillantes en train de former des étoiles.

Dans le cadre du sondage *Canada - France Deep Field*, [Foucaud et al. \(2003\)](#) observent 1294 LBGs à $z \sim 3$. Une longueur de corrélation de $r_0 = 5.9 \pm 0.5 \text{ h}^{-1} \text{ Mpc}$ est ainsi mesurée pour l'ensemble de leurs 3 champs à partir de la fonction de corrélation angulaire. Ils en dérivent un biais linéaire $b = 3.5 \pm 0.3$ à partir de prédictions du champ de densité de matière noire. De plus, une ségrégation en couleur et luminosité est visible. Comme dans l'Univers local (partie 1.5.1), les LBGs les plus rouges, tout comme les plus brillantes, présentent une amplitude de corrélation plus forte que celles des bleues et faibles respectivement (voir aussi [Giavalisco et Dickinson \(2001\)](#)).

Une sélection dans un diagramme couleur permet aussi de considérer des galaxies de *redshift* compris entre 1.4 et 3.5. Comme précédemment mais avec 28500 galaxies réparties dans 21 champs (surface totale de 0.81 degré carré), [Adelberger et al. \(2005\)](#) mesurent la longueur de corrélation spatiale r_0 à 3 *redshifts* moyens (table 1.5) à partir de $\omega(\theta)$ et d'une distribution en *redshifts* établie à partir de 1600 spectres.

TAB. 1.5 – Longueur de corrélation r_0 (exprimée en $\text{h}^{-1} \text{ Mpc}$) de LBGs ([Adelberger et al., 2005](#))

\bar{z}	r_0	masse du halo (M_\odot) (log(m))
1.7	4.5 ± 0.6	11.9 – 12.3
2.2	4.2 ± 0.5	11.8 – 12.5
2.9	4.0 ± 0.6	11.2 – 11.8

Des observations utilisant la même technique ont été réalisées à $z \sim 4$ par [Ouchi et al. \(2001\)](#) et jusqu'à des *redshifts* de ~ 5 par [Ouchi et al. \(2004\)](#), [Allen et al. \(2005\)](#) et [Lee et al. \(2005\)](#). Il a été possible d'atteindre de tels décalages spectraux grâce à des données obtenues avec le télescope spatial Hubble combinées avec des observations au sol sur les grands télescopes de classe 8-10m (Subaru et Keck). Une compilation de leurs mesures ainsi que d'autres à $z \sim 3$ sont reportées figure 1.22.

On retiendra que les galaxies étudiées à très grands *redshifts* grâce à la technique de la cassure de Lyman à 912\AA sont des traceurs fortement biaisés du champ de densité de la masse. De plus, une ségrégation en luminosité est déjà observée dans ces popula-

tions. Les plus lumineuses sont les plus agrégées. Elles se trouvent enfin dans des halos de matière noire très massifs et correspondraient aux progéniteurs de galaxies massives observées localement.

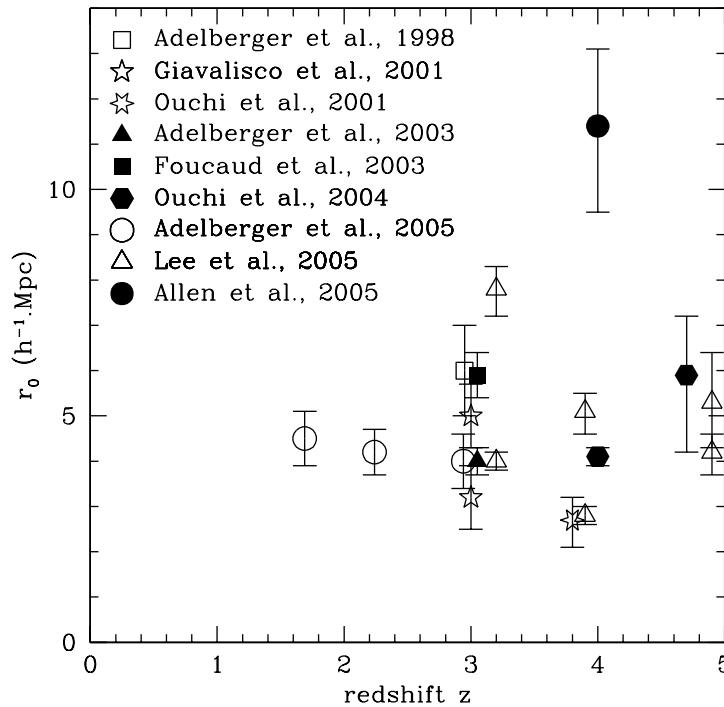


FIG. 1.22 – Longueur de corrélation de population de galaxies de type *Lyman-Break* identifiées à grands *redshifts*.

1.5.2.3 Les galaxies extrêmement rouges

Dans le cadre d'un sondage de galaxies sélectionnées en bande K, [Carlberg et al. \(1997\)](#) observent que leurs 248 objets sont plus fortement corrélés que ceux sélectionnés traditionnellement en optique. L'amplitude de la fonction de corrélation ne semble pas évoluer avec le *redshift* mais le manque de statistique ne permet pas de conclure. Cette dernière est améliorée par [Daddi et al. \(2000\)](#) qui observent un échantillon complet de 398 EROs (définition partie 1.4.2.5) avec $(R-K_s) \geq 5$ dans un champ de 701 minutes d'arc carrés. La distribution spatiale de ces objets est très inhomogène. Ils mesurent une forte amplitude de la fonction de corrélation angulaire qu'ils interprètent comme le signal de galaxies elliptiques à $z \sim 1$. Une modélisation de la distribution en *redshifts* de ces galaxies EROs permet de quantifier la longueur de corrélation spatiale. [Daddi et al. \(2001\)](#) mesurent ainsi $r_0 = 12 \pm 3 \text{ h}^{-1} \text{ Mpc}$ à $K_s = 19.2$. De leur côté, [Firth et al. \(2002\)](#) mesurent des longueurs de corrélation spatiales du même ordre de grandeur comprises entre ~ 7.5 et $\sim 10.5 \text{ h}^{-1} \text{ Mpc}$ pour un échantillon de galaxies EROs sélectionnées par leur couleur $(I-H) > 3$ (pour un intervalle en *redshift* $[0.9 - 1.5]$).

Comparés aux mesures faites par [Guzzo et al. \(1997\)](#) et [Willmer et al. \(1998\)](#), ces

résultats impliqueraient une faible évolution de l'amplitude de l'agrégation des elliptiques brillantes entre $z \sim 1-1.5$ et $z \sim 0$.

Dans le cadre d'un modèle de conservation des galaxies où les elliptiques sont formées à grands *redshifts* et ne font que suivre l'évolution des structures, les prédictions pour le biais et la longueur de corrélation spatiale permettent de relier les galaxies EROs aux LBGs (section 1.5.2.2). Mais l'amplitude de $\xi(r)$ mesurée par Daddi *et al.* (2001) ne semble pas diminuer assez entre $z \sim 0$ et $z \sim 1$ pour aller dans ce sens. Un modèle basé sur la fusion des galaxies (Mo et White, 1996) dans lequel le biais augmente avec le *redshift* semble lui plus en accord avec la forte agrégation des galaxies EROs.

1.6 Évolution des grandes structures : prévision des modèles

Au cours de la fin du 20^{ième}, début du 21^{ième} siècle, le développement de l'informatique et des supers calculateurs a permis de nombreuses avancées dans l'astrophysique numérique. Cette dernière consiste à produire des simulations toujours plus réalistes et avec toujours plus de résolution. Ces simulations sont ensuite utilisées pour faire des prédictions et être confrontées aux observations.

Les premiers travaux sur les simulations à N-corps souffraient de leur faible résolution, rendant difficile le sondage de grands volumes tout en étant capable de résoudre les halos de matière noire contenant les galaxies à étudier. Pour localiser les galaxies parmi l'ensemble des particules de matière noire utilisées, des prescriptions très simples étaient alors employées, en ajustant par exemple la densité moyenne des galaxies, généralement fonction de la densité locale, lissée sur des échelles de plusieurs méga parsecs (Cole *et al.*, 1998; Yoshida *et al.*, 2001). Puis des approches plus fines furent utilisées dans des simulations à grande résolution. Les particules de matière sombre furent considérées comme des éléments d'un halo (Kauffmann *et al.*, 1997). L'idée importante est alors de dire que les galaxies se forment dans des halos virialisés et donc que leurs propriétés dépendent en grande partie de la masse du halo.

Ces simulations numériques permettent de tester les modèles cosmologiques et de faire des prédictions pour la fonction de corrélation spatiale, la fonction de luminosité, les dépendances à la luminosité, à l'échelle sondée, à l'environnement, ... Les observations permettent de vérifier, d'ajuster, ou de modifier les hypothèses et conditions initiales choisies pour chaque simulation.

Les simulations actuelles de matière noire sont de plus en plus réalistes (Springel *et al.*, 2005). Par contre il est impossible de les confronter directement aux observations car elles ne manipulent que la matière noire. Il faut alors rajouter des modèles (semi-)analytiques de formation et d'évolution des galaxies. Ces modèles sont généralement réglés de manière à reproduire correctement les observations dans l'Univers local. Des prédictions peuvent ensuite être faites à grands *redshifts*.

L'utilisation de la très grande simulation du *Hubble-Volume* (voir partie 4.1.1) a tout d'abord permis une étude de l'agrégation des halos de matière noire pour quatre modèles cosmologiques différents et de comparer ces résultats aux mesures du sondage APM (partie 1.3.3). Jenkins *et al.* (1998) concluent que quelque soit le modèle, la distribution des galaxies est soit biaisée soit anti-biaisée par rapport à celle de la masse sous-jacente.

De plus, l'expression de ce biais n'est pas simple et dépend de l'échelle considérée.

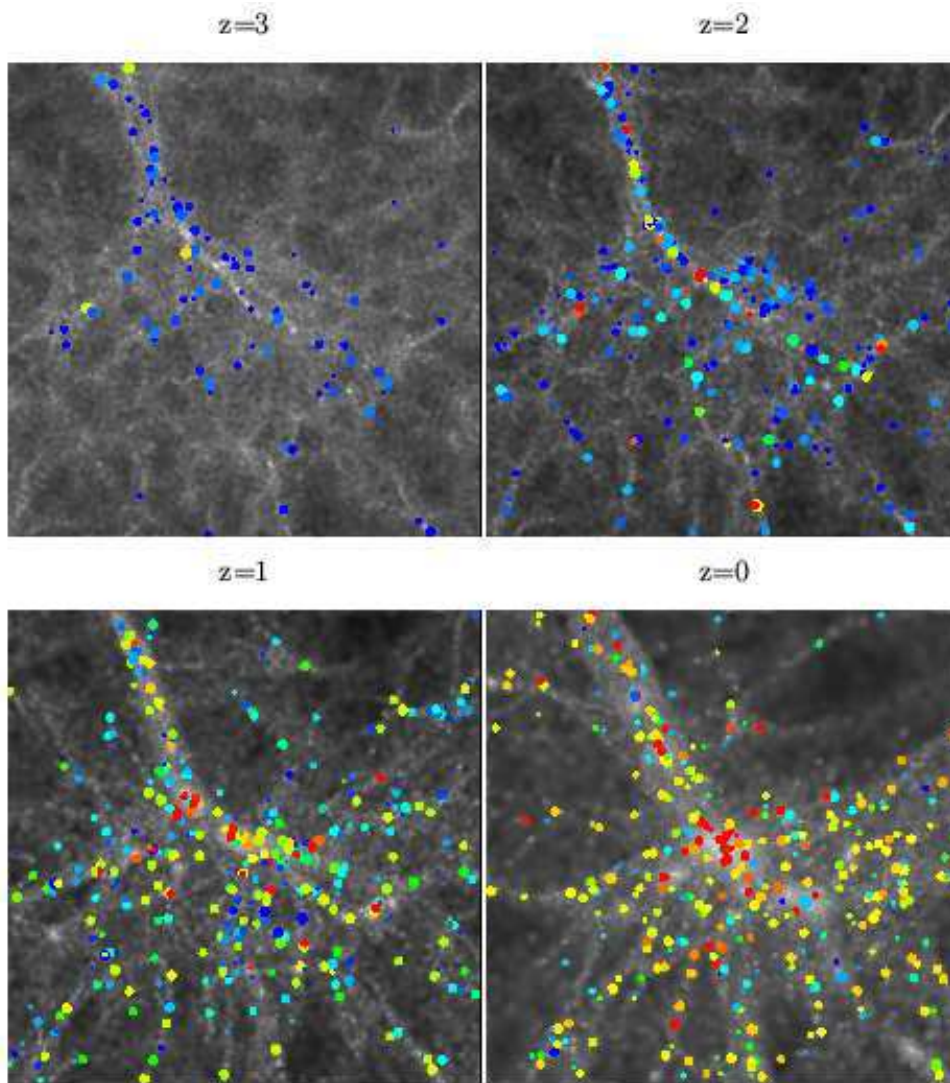


FIG. 1.23 – Distribution spatiale des galaxies (couleur) superposées à la distribution de matière noire à partir des simulations de la collaboration GIF (Kauffmann *et al.*, 1999a). La couleur des galaxies, du bleu au rouge est fonction du taux de formation d'étoiles. Les quatre panneaux montrent l'évolution de la distribution des galaxies et de la matière noire de $z=3$ à $z=0$.

Couplées à des modèles semi-analytiques afin de décrire les galaxies, ces simulations ont permis de réaliser de nombreuses prédictions quant à l'évolution de l'agrégation des galaxies aux grandes échelles. La figure 1.23 montre l'évolution de la distribution spatiale des galaxies et de la matière noire entre $z=3$ et $z=0$. L'agrégation des galaxies dépend de la luminosité, de la couleur, du type, de la morphologie ou encore du taux de formation d'étoiles des galaxies identifiées (Kauffmann *et al.*, 1999a,b; Weinberg *et al.*, 2004). Benson *et al.* (2001) prédisent aussi qu'une relation morphologie-densité (Dressler, 1980) est déjà en place au moins à *redshift* $z=1$.

La figure 1.24 représente dans le cadre de deux modèles cosmologiques des prédictions de la longueur de corrélation en fonction du *redshift* selon que les galaxies soient sélectionnées en bande B, en bande I, selon qu'elles soient brillantes, qu'elles aient un fort taux de formation stellaire ou encore qu'elles soient de type précoces (Kauffmann *et al.*, 1999b). Ainsi par exemple, l'amplitude des galaxies précoces est beaucoup plus importante que la population globale et varie peu en fonction du *redshift*. Dans un modèle τ CDM, l'évolution est moins sensible à l'échantillon considéré : r_0 reste stable jusqu'à $z \sim 1$ puis augmente rapidement.

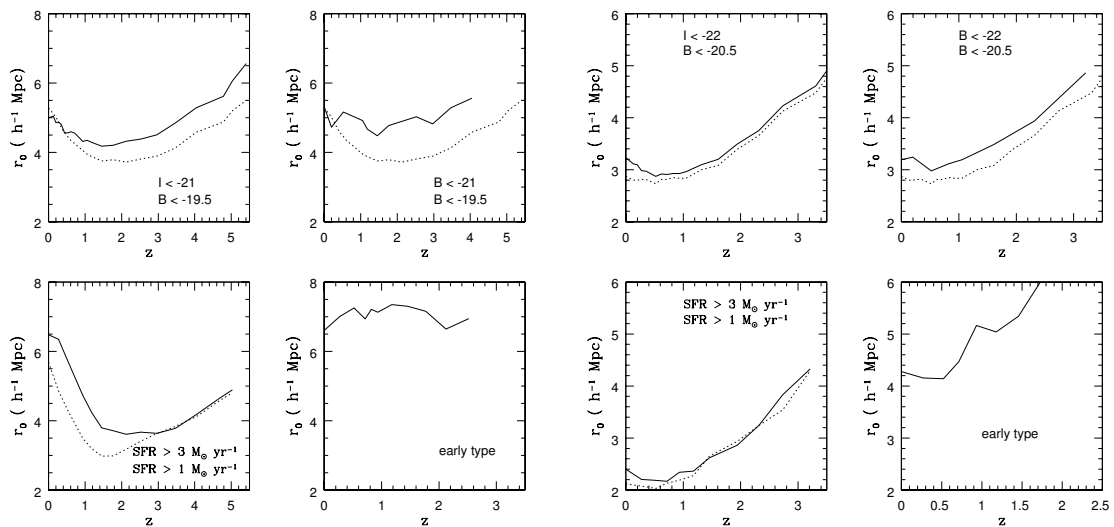


FIG. 1.24 – Évolution de la longueur de corrélation spatiale dans un modèle Λ CDM (les 4 panneaux de gauche) et dans un modèle τ CDM (les 4 de droite). Pour chaque cosmologie, les graphiques (a), (b), (c) et (d) sont répartis de haut en bas et de gauche à droite avec : (a) galaxies L_* sélectionnées en bande I et B au repos, (b) galaxies brillantes et de type L_* , (c) galaxies sélectionnées en fonction de leur taux de formation stellaire, (d) galaxies précoces de masse supérieure à $3 \times 10^{10} M_\odot$ (Kauffmann *et al.*, 1999b, figures 7 et 8).

La figure 1.25 indique l'évolution de la longueur de corrélation r_0 et de la pente γ de la fonction de corrélation spatiale en fonction de la masse des galaxies, comparée à celle de la matière noire (Weinberg *et al.*, 2004). L'évolution de la longueur de corrélation des galaxies est lente mais dépend significativement de leur masse et de leur âge. Des précautions doivent donc être prise dans l'interprétation de résultats de sondages qui observent des populations de différentes masses à différents *redshifts*.

La figure 1.25 met aussi en évidence que le biais existant entre la distribution des galaxies et celle de la matière noire augmente avec le *redshift*. De nombreux modèles de biais ont été développés. On retiendra trois approches différentes (Moscardini *et al.*, 1998) :

- Un modèle où le nombre de galaxies se conserve en fonction du temps (Fry, 1996). Dans ce schéma, une galaxie formée à un *redshift* z_f avec un biais b_f ne fait que suivre l'évolution des structures sous l'effet de la gravité. Tegmark et Peebles (1998) donnent une description de l'évolution du biais pour ce scénario.

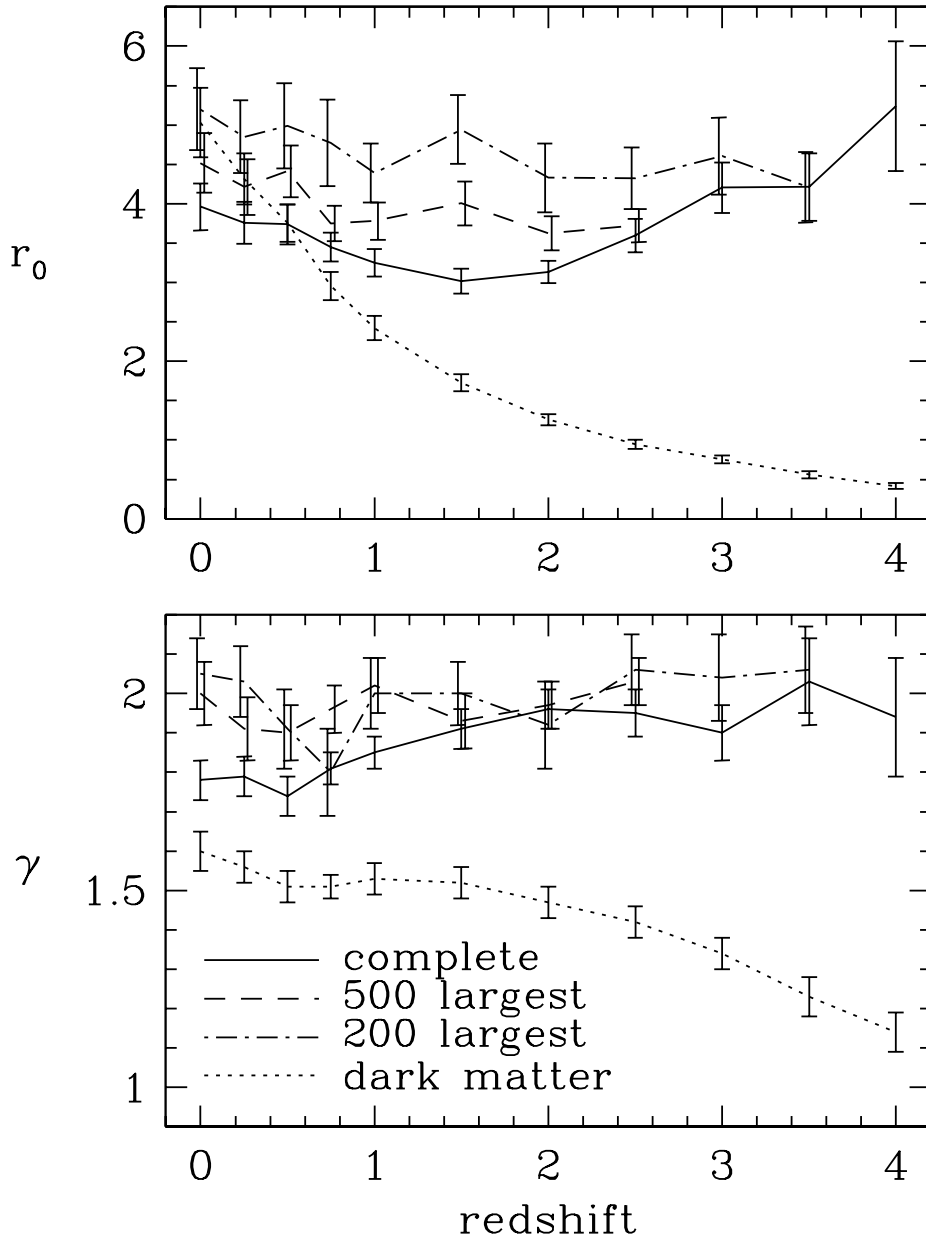


FIG. 1.25 – Évolution de la longueur de corrélation spatiale r_0 (en $h^{-1} \text{Mpc}$) et pente γ pour toutes les galaxies (ligne continue), pour les 500 plus massives (tirets), pour les 200 plus massives (point-tirets) et pour la matière noire (pointillés). Les courbes d'évolution des 500 plus massives galaxies s'arrêtent à $z=2.5$ car l'échantillon complet n'en contient pas suffisamment au-delà. Pour la même raison, les courbes relatives à l'échantillon des 200 plus massives galaxies s'arrêtent à $z=3.5$. (Weinberg *et al.*, 2004, figure 5)

- Un modèle où l'on prend en compte la fusion des galaxies. [Mo et White \(1996\)](#) fournissent ainsi des prédictions analytiques pour le calcul du biais.
- Un modèle où la distribution des galaxies plus brillantes qu'une luminosité L trace bien la distribution de halos de masses supérieures à M . Ce modèle prédit le biais d'objets qui viennent juste de se former au *redshift* d'observation ([Blanton et al., 2000](#))

Ces modèles, établis à partir d'approximations de la complexité des phénomènes physiques mis en jeu, permettent de réaliser des prédictions qui seront ensuite comparées aux observations et/ou simulations numériques. Ainsi, à partir d'une simulation à N corps couplée à des recettes semi-analytiques, [Somerville et al. \(2001\)](#) constatent qu'un modèle dans lequel le nombre de galaxies est conservé sous-estime l'évolution du biais alors qu'un modèle qui tient compte de la fusion des galaxies en fournit une bonne description.

L'une des difficultés des simulations est de bien décrire les galaxies au sein des halos de matière noire. Une approche simple consiste à ne placer qu'une seule galaxie par halo. Une description plus détaillée et très prometteuse, dite HOD (pour *Halo Occupation Distribution*), décrit la distribution des galaxies à l'intérieur de chaque halo, et donc le biais, en définissant la probabilité $P(N|\langle N \rangle)$ qu'un halo virialisé de masse M contienne N galaxies d'une certaine classe. Le nombre moyen de galaxies par halo de matière noire de masse M peut par exemple être défini comme une loi de puissance au-delà d'une masse minimum M_{\min} ([Berlind et Weinberg, 2002](#)) :

$$\langle N \rangle = \begin{cases} 0 & \text{si } M < M_{\min} \\ (M/M_1)^\alpha & \text{si non} \end{cases} \quad (1.25)$$

La masse M_1 correspond à la masse d'un halo qui contient en moyenne une seule galaxie. La forme la plus simple conduit à $\langle N \rangle$ proportionnel à M et donc $\alpha = 1$. Mais les différents processus physiques affectent la valeur de la pente comme par exemple les fusions de galaxies ou leur évolution morphologique. Cette description de la façon dont les halos de matière noire sont peuplés apporte une compréhension des structures à grande échelle dans le régime non-linéaire. De nombreux modèles semi-analytiques implémentent maintenant cette approche pour modéliser les halos ([Kauffmann et al., 1999a](#); [Somerville et Primack, 1999](#); [Colberg et al., 2000](#); [Benson et al., 2001](#); [Berlind et al., 2003](#)). Ils permettent de faire de nombreuses prédictions et d'apporter des contraintes aux modèles de formation des galaxies.

1.7 Synthèse

Les grands sondages de galaxies, combinant photométrie et spectroscopie, ont mis en évidence la présence de grandes structures dans l'Univers, aussi bien localement qu'à des époques où l'Univers n'était âgé que de 2 à 3 milliards d'années. Dans le modèle dit *standard*, ces structures sont le résultat de l'évolution des fluctuations primordiales du champ de densité de matière noire. Les petites structures se sont formées les premières, et leurs fusions ont abouti à la création de structures plus importantes (modèle *hiérarchique*).

Les modèles théoriques permettent maintenant de relativement bien décrire l'évolution des halos de matière noire dans lesquels se sont formées les galaxies ainsi que de faire des prédictions. Les galaxies sont en fait un traceur biaisé du champ de densité et ce biais

dépend des propriétés physiques des galaxies. Ainsi, leur agrégation dépend fortement de leur luminosité, de leur couleur, de leur type spectral et morphologique. Les galaxies lumineuses sont ainsi plus corrélées que les faibles. L'amplitude de la fonction de corrélation des galaxies bleues est plus importante que celle des rouges tout comme l'est la longueur de corrélation des galaxies de type précoce (elliptiques, lenticulaires) par rapport à celle des galaxies tardives (spirales, irrégulières).

Ces résultats sont maintenant bien établis dans l'Univers proche comme nous l'avons montré dans ce chapitre, mais peu d'observations facilement comparables pour retracer l'évolution et dans des champs suffisamment grands pour s'affranchir de la variance cosmique n'avaient été réalisées jusqu'au début de cette thèse en 2002. Pourtant il est important de pouvoir observer un nombre statistiquement représentatif de galaxies à différentes époques, jusqu'à de très grands *redshifts* afin de pouvoir suivre l'évolution et la façon dont se sont formées les galaxies et les grandes structures. C'est un des buts du sondage profond *VIMOS-VLT Deep Survey* présenté dans le chapitre suivant.

Chapitre 2

Le sondage profond VIMOS-VLT *Deep Survey*

Résumé

2.1 Des sondages d'Univers lointain	49
2.2 Le sondage spectroscopique profond VVDS	51
2.3 Réduction des données spectroscopiques et détermination des <i>red-shifts</i>	58
2.4 Première campagne d'observation : les données	73

Ce chapitre présente le sondage spectroscopique VIMOS-VLT Deep Survey. Je décris les caractéristiques de ce sondage profond de l'Univers. Je présente ensuite comment ont été réduites les données spectroscopiques et comment les redshifts ont été mesurés. Je décris enfin les données de la première campagne d'observation qui ont été utilisées par la suite.

2.1 Des sondages d'Univers lointain

Mon travail de thèse s'est déroulé dans le cadre du sondage spectroscopique profond du VVDS. Avant d'en faire sa description dans la suite de ce chapitre, je décris ici les principaux sondages dont le but est d'observer l'Univers distant et les compare au VVDS.

Comme nous l'avons dit dans le premier chapitre, un sondage du ciel doit savoir faire le bon compromis entre couverture du ciel, profondeur atteinte, résolution spectroscopique et échantillonnage spatiale. Le tableau 2.1 décrit quelques éléments clefs de plusieurs de ces sondages : la taille du champ, la profondeur atteinte dans une certaine bande, le nombre d'objets (photométriques et/ou spectroscopiques), et le nombre de bandes photométriques utilisées.

Le projet américain DEEP2 ([Davis et al., 2003](#)) consiste à obtenir le spectre de quelques 55000 galaxies à $z > 0.7$. Ces observations sont réalisées à partir du spectrographe multi-objets à fentes DEIMOS, installé sur le télescope Keck II de 10 mètres de diamètre situé au sommet du Mauna Kea à 4205 mètres d'altitude. Jusqu'à 75 spectres de résolution

$R \sim 4000$ peuvent être enregistrés simultanément. La résolution de ces spectres permet par exemple d'établir les courbes de rotations des galaxies. Les cibles sont sélectionnées selon un critère en couleur $R-I$ et $B-R$ (Coil *et al.*, 2004). L'avantage de cette technique est de ne conserver effectivement que des galaxies à $z \sim 1$. Par contre il est difficile d'estimer le nombre et les propriétés des objets écartés mais pourtant situés à ces grands *redshifts*. Cela complique alors leur objectif scientifique qui est d'étudier les propriétés et l'agrégation des galaxies à $z \sim 1$ et de les comparer aux mesures réalisées dans l'Univers local par les sondages 2dFGRS et SDSS.

TAB. 2.1 – Paramètres important qui conditionnent l'efficacité d'un sondage

sondage	taille (en deg ²)	profondeur	nombre d'objets	nombre de filtres
VVDS <i>Wide</i>	16	$I_{AB} \leq 22.5$	$\sim 10^5$ ^j	BVRI
VVDS <i>Deep</i>	1.2	$I_{AB} \leq 24$	~ 30000 ^j	UBVRIJK
DEEP2 ^a	3.5	$R_{AB} \leq 24.1$	~ 55000 ^j	BRI
COSMOS	2	$I_{AB} \leq 27.5$	$\sim 2 \cdot 10^6$	~ 20 ^b
GEMS ^c	0.22	$V_{AB} \leq 28.3$	~ 10000	F606W, F850LP
COMBO17 ^d	1 (4 champs)	$R_{AB} \leq 24$	~ 60000 ⁱ	17 dont UBVRi
CNOC2 ^e	1.5 (4 champs)	$R_c \leq 21.5$	~ 6000 ^j	UBVR _c I _c
CFRS ^f	0.14 (5 champs)	$I_{AB} \leq 22.5$	591 ^j	BVIK
CFDF	4×0.24	$\sim 26-27$	~ 100000 ⁱ	UBVI
2dFGRS ^g	2000	$B_J \leq 19.45$	$\sim 2.3 \times 10^5$ ^j	b _J
SDSS ^h	6670 ⁱ 4783 ^j	$r=22.2$	$\sim 1.8 \times 10^8$ ⁱ $\sim 8.5 \times 10^5$ ^j	ugriz

^aDavis *et al.* (2003)

^bHST, CFHTLS, Subaru, Spitzer, radio, XMM, ...

^c(Rix *et al.*, 2004)

^dWolf *et al.* (2004)

^eYee *et al.* (2000)

^fLilly *et al.* (1995a)

^gColless *et al.* (2001)

^hAdelman-McCarthy (2005)

ⁱimagerie

^jspectroscopie

La surface du champ couverte par DEEP2 est relativement grande comparée à des sondages comme le CNOC2, le CFRS ou même COMBO17. Ces champs là souffrent des effets de la variance cosmique. Le projet COMBO17 (Wolf *et al.*, 2004) essaie lui de s'en affranchir en étant capable, par une technique de *redshifts* photométriques, d'obtenir l'information à trois dimensions de plus de 60000 objets, soit autant que pour DEEP2 mais pour un champ 4 fois plus petit. L'inconvénient majeur reste la faible précision des *redshifts* photométriques ($\delta z / (1+z) \sim 0.02$). Le projet GEMS (Rix *et al.*, 2004, *Galaxy Evolution from Morphologies and SEDs*) combine lui les informations déjà disponibles sur le champ CDFS grâce aux sondages GOODS (Giavalisco *et al.*, 2004) et COMBO17 en l'élargissant à $28' \times 28'$ à partir d'observations réalisées dans deux couleurs avec le télescope spatial Hubble. Ce nouveau sondage permet d'étudier ~ 10000 galaxies en fonction de leur morphologie, ce que ne peuvent pas réaliser les autres sondages d'Univers distants. Un

autre sondage en cours, le projet COSMOS, fournira, là aussi grâce à des images HST, l'information morphologique de $\sim 10^6$ galaxies sur un champ encore 4 fois plus grand.

De tous ces sondages, seuls le sondage COSMOS et le VVDS observent des volumes d'Univers comparables à ceux du SDSS et 2dFGRS. Les galaxies de ce dernier remplissent un volume d'Univers de $\sim 10^8 h^{-3} \text{Mpc}^3$ jusqu'à $z \sim 0.3$ (Colless *et al.*, 2001), à comparer aux $\sim 9 \times 10^7 h^{-3} \text{Mpc}^3$ du sondage COSMOS jusqu'à $z \sim 4$ ou encore aux $\sim 1.5 \times 10^6 h^{-3} \text{Mpc}^3$ du VVDS jusqu'à $z \sim 1.5$ (Marinoni *et al.*, 2005). Le sondage VVDS est le pendant à grand *redshift* des grands sondages de l'Univers local.

2.2 Le sondage spectroscopique profond VVDS

2.2.1 But et stratégie

Le VIMOS-VLT Deep Survey¹ est un très grand relevé profond du ciel. Le projet, lancé en 1996, est conduit par un consortium franco-italien de 47 chercheurs, post-doctorants et doctorants. Il regroupe le Laboratoire d'Astrophysique de Marseille, l'Istituto di Fisica Cosmica e Tecnologia Relative (Milan), l'Osservatorio Astronomico di Brera (Milan), l'Osservatorio Astronomico di Bologna (Bologne), l'Osservatorio Midi-Pyrénées, l'Osservatorio Astronomico di Capodimonte (Naples), et l'Institut d'Astrophysique de Paris. Ce consortium a été chargé de la construction du spectrographe multi-objets VIMOS (voir partie 2.2.2) en échange d'une utilisation garantie de l'instrument à hauteur de 50 nuits d'observation.

Le but du VVDS est d'étudier la formation et l'évolution des galaxies, celle des grandes structures, des amas de galaxies ainsi que les galaxies à noyau actif en remontant jusqu'à une époque où l'Univers était âgé de moins de 10% de son âge actuel.

Les observations du ciel sont menées sur 4 grands champs indépendants situés sur l'équateur céleste (figure 2.1) : 0226-0430, 1003+0154, 1400+0500 et 2217+0024. Chaque champ mesure 2×2 degrés carrés. Ceci permet de tenir compte de la variance cosmique et d'étudier les grandes structures jusqu'à des échelles de $100 h^{-1} \text{Mpc}$ à *redshift* ~ 1 . Enfin, un cinquième champ a été rajouté à cette liste : le CDFS (pour *Chandra Deep Field South*).

Le relevé se divise en deux étapes successives : l'observation profonde des quatre champs en imagerie suivie d'observations spectroscopiques.

La première étape consiste à réaliser un relevé photométrique des champs définis ci-dessus dans différents filtres. Ces observations se sont déroulées en grande partie au télescope Canada-France-Hawaii avec la caméra CFH12k au travers des filtres B, V, R et I sur les quatre champs (Le Fèvre *et al.*, 2004a; McCracken *et al.*, 2003). Des observations complémentaires en bande U ont été obtenues sur le télescope de 2.2m de l'ESO à La Silla (Radovich *et al.*, 2004) et en bande J et K sur le NTT (Iovino *et al.*, 2005) mais ne couvrent pas tous ces champs. Le tableau 2.2 résume les observations obtenues pour ces quatre champs dans les différents filtres (leurs courbes de transmission sont reportées figure 2.2) ainsi que les données EIS (Arnouts *et al.*, 2001) utilisées pour le CDFS.

Le traitement des données photométriques s'est déroulé au sein de l'équipe TERAPIX² de l'Institut d'Astrophysique de Paris, à l'exception des données infrarouges réduites par

¹voir <http://www.oamp.fr/virmos/>

²voir <http://terapix.iap.fr/>

Champ	R.A. (J2000)	Déc. (J2000)	Filtre	Magnitude de complétude (AB)	Surface (deg ²)
VVDS-02h	02 : 26 : 00	-04 : 30 : 00	U	25.0	0.71
			B	26.5	1.2
			V	26.2	1.2
			R	25.9	1.2
			I	25.0	1.2
			J	22.25	$4.5 \cdot 10^{-2}$
			K	20.75	$4.6 \cdot 10^{-2}$
			GALEX/NUV	24.4	1.2
			GALEX/FUV	24.8	1.2
			Spitzer/3.6 μ m	20.4	1.1
			Spitzer/4.5 μ m	19.1	1.1
			Spitzer/5.8 μ m	15.7	1.1
			Spitzer/8 μ m	15.9	1.1
			Spitzer/24 μ m	11.9	1.1
VVDS-10h	10 : 03 : 39	+01 : 54 : 39	I	25.0	4
			J	22.25	$2.9 \cdot 10^{-2}$
			K	20.75	$3.4 \cdot 10^{-2}$
VVDS-14h	14 : 00 : 00	+05 : 00 : 00	I	25.0	4
			J	22.0	$6.5 \cdot 10^{-3}$
			K	20.75	$1.3 \cdot 10^{-2}$
VVDS-22h	22 : 17 : 50	+00 : 24 : 27	I	25.0	4
			J	21.9	$2.8 \cdot 10^{-2}$
			K	21.0	$2.5 \cdot 10^{-2}$
VVDS-CDFS	03 : 32 : 28	-27 : 48 : 30	U	25.7	0.25
			U'	26.0	0.25
			B	26.4	0.25
			V	25.4	0.25
			R	25.5	0.25
			I	24.7	0.25

TAB. 2.2 – Résumé des données photométriques disponibles pour les différents champs du VVDS. Les magnitudes de complétude sont données à 50% pour les données CFH12k (dans une ouverture de 3 secondes d'arc), et à 5σ pour les données GALEX ([Morrissey et al., 2005](#)), Spitzer, et du CDFS. Pour des informations complémentaires, se référer à [McCracken et al. \(2003\)](#), [Radovich et al. \(2004\)](#), [Iovino et al. \(2005\)](#) et [Arnouts et al. \(2001\)](#).

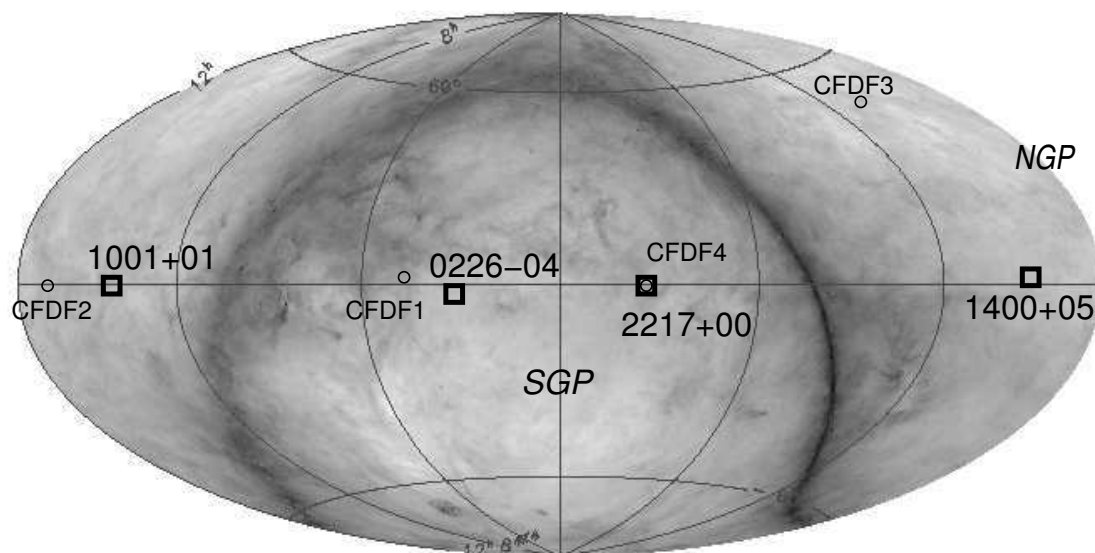


FIG. 2.1 – Position des quatre champs VVDS sur une carte d'extinction du ciel (Le Fèvre *et al.* (2004a)). Les carrés représentent la position des 4 champs larges.

l'équipe de Milan-Brera. Ainsi, un ensemble de plus de 2 millions d'objets extraits de ces images sont stockés dans une base de données CENCOS du Laboratoire d'Astrophysique de Marseille³. Elle contient entre autre les coordonnées et magnitudes dans chacune des bandes disponibles pour chaque objet du catalogue.

La deuxième partie du sondage consiste à déterminer avec VIMOS le *redshift* de plus de 100000 galaxies extraites du catalogue de la base de donnée. Le critère de sélection des objets pour la partie spectroscopie est le plus simple possible afin d'introduire le moins de biais observationnels possible. Il consiste à choisir au hasard des objets parmi ceux plus brillants qu'une certaine limite. Contrairement à d'autres sondages de galaxies (DEEP2 (Davis *et al.*, 2003), Lyman-Break (Steidel *et al.*, 1996), ...) qui combinent une sélection en magnitude à des critères de couleur à partir de plusieurs bandes photométriques, une pure sélection en magnitude ne sélectionnera pas a priori telle ou telle classe d'objet. L'échantillon alors obtenu sera donc représentatif de la population globale, ce qui est le but recherché.

La bande de sélection choisie est la bande I qui couvre la plage en longueur d'onde entre 700 et 950 nanomètres. Le choix de la bande I réalise un compromis entre la sélection de galaxies à partir du flux émis par leurs étoiles vieilles au-dessus de 4000Å jusqu'à $z \sim 1$, et la sélection de galaxies détectées par leur continu émis en-dessous de 4000Å jusqu'à l'ultraviolet pour les objets à *redshift* au delà de 1. Là encore, la partie spectroscopie du sondage est divisée en deux. La première partie dite *Wide* consiste à enregistrer le spectre et à mesurer le *redshift* d'objets détectés dans la bande I avec une magnitude (AB) comprise entre 17.5 et 22.5 sur les quatre champs du VVDS alors que la deuxième partie, *Deep*, est plus profonde jusqu'à $I_{AB} = 24$, mais sur seulement 1.2 degrés carrés du champ VVDS-02h et une partie du champ VVDS-CDFS. C'est la première fois qu'un

³voir <http://cencosw.oamp.fr/>

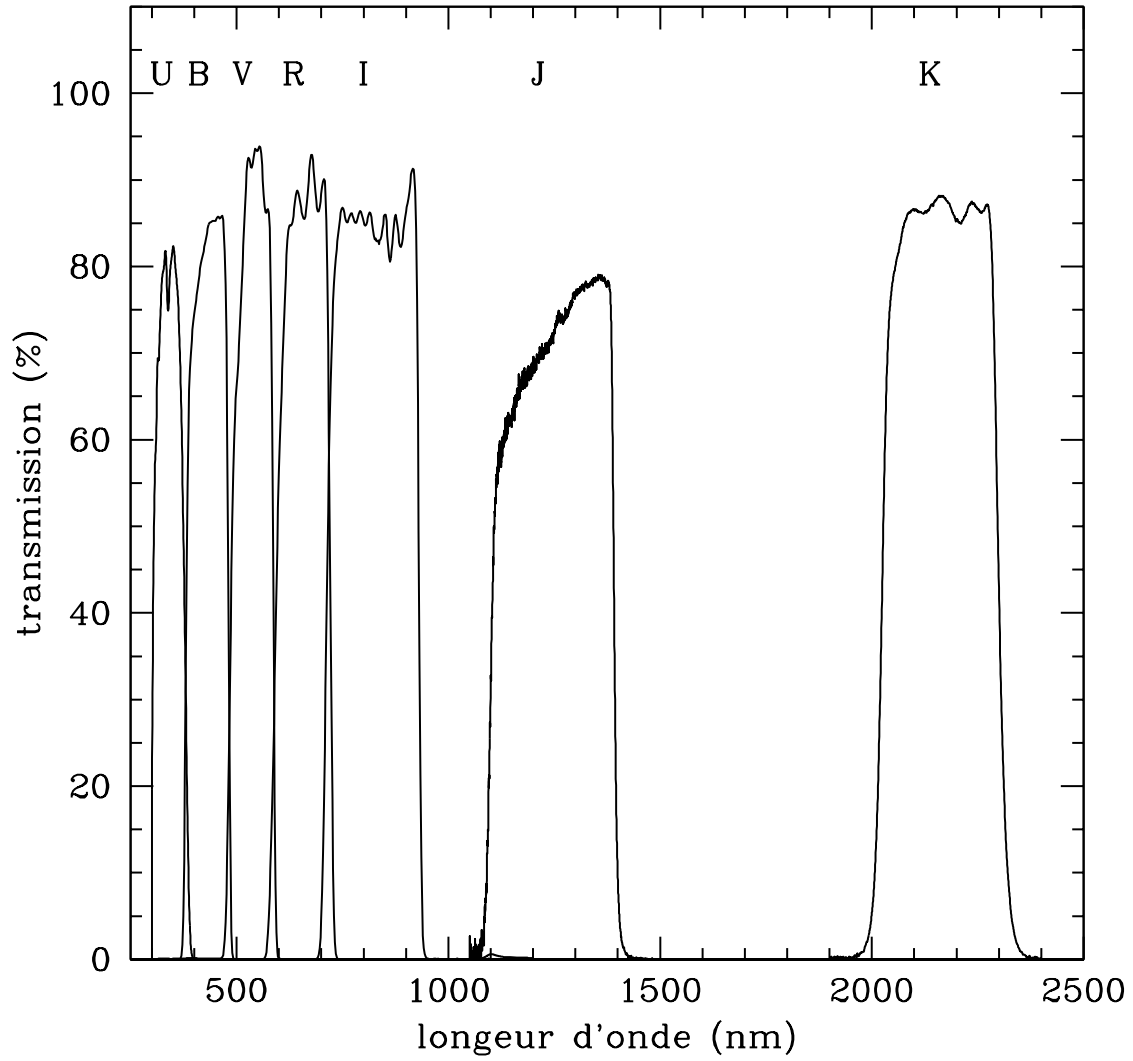


FIG. 2.2 – Courbe de transmission des filtres B, V, R et I de la caméra CFH12k/CFHT, U de la caméra WFI/ESO-MPG, J et K de la caméra SOFI/ESO-NTT

sondage complet en magnitude est conduit à une magnitude aussi profonde de $I_{AB} = 24$, sur un échantillon aussi grand. La partie *Wide* permet d'étudier les grandes structures et d'évaluer la variance cosmique alors que la partie *Deep* sonde les galaxies à plus grand *redshift* et apporte des informations sur les objets les plus faibles.

2.2.2 L'instrument VIMOS

Le large programme VVDS est rendu possible par l'utilisation du spectrographe multi-objets VIMOS (pour *Visible Multi-Object Spectrograph*). Cet instrument est installé au foyer Nasmyth de l'unité Melipal du *Very Large Telescope* (VLT) de l'Observatoire Austral Européen (ESO), situé au sommet du mont Paranal à 2635 mètres d'altitude dans la cordillère des Andes. Il exploite au maximum les capacités d'acquisition du télescope de 8 mètres de diamètre ce qui permet de réaliser des sondages spectroscopiques profonds. VIMOS est capable d'enregistrer simultanément jusqu'à 1000 spectres de galaxies (600 en moyenne) en basse résolution ($R \sim 200$) (Le Fèvre *et al.*, 2003) et rend donc possible la réalisation de grands sondages. À titre de comparaison, un relevé de la taille du Canada France *Redshift* Survey (CFRS, voir Le Fèvre *et al.* (1995)) qui avait nécessité plus de 25 nuits d'observation avec MOS-SIS au télescope CFHT il y a seulement une dizaine d'années peut ainsi être mené à bien en une heure et demie de pose,

VIMOS, développé et construit pour le compte de l'ESO par le consortium franco-italien VVDS, a été testé et calibré à l'Observatoire de Haute Provence. Il y a reçu sa première lumière dite "technique" le 12 mai 2000. Démonté puis réintégré au Chili, il est opérationnel depuis le 26 février 2002. Il peut fonctionner dans trois modes : imagerie, spectroscopie multi-objet, spectroscopie intégrale de champ.

2.2.2.1 Imagerie et spectroscopie multi-objet

Le champ de vue de VIMOS se sépare en 4 canaux (ou quadrants) chacun de 7 par 8 minutes d'arc environ séparés par un motif en croix de largeur de 2 minutes d'arc (figure 2.3). Chaque canal est indépendant et représente un spectrographe multi-objet à part entière. La chaîne d'acquisition contient un filtre interférentiel afin de choisir la bande passante en longueur d'onde. Dix filtres sont disponibles. Un masque usiné sur place grâce à l'unité de fabrication *Mask Manufacturing Unit* (MMU, Conti *et al.* (2001)) peut être inséré. Il contient un certain nombre de fentes parallèles entre elles et déterminées afin de ne sélectionner que les objets (galaxies/étoiles) que l'on cherche à étudier. Un réseau disperse ensuite leur lumière perpendiculairement à la direction des fentes. Le choix du réseau se fait parmi six d'entre eux selon la résolution en longueur d'onde ($\lambda/\Delta\lambda$ compris entre 200 et 2500, pour une fente de 1 seconde d'arc) et la plage du spectre voulue (3700Å - 6700Å pour le grisme bleu et 5000Å - 10000Å pour le grisme rouge). L'image des spectres est enfin enregistrée sur un capteur CCD de 2048×4096 pixels. Sa résolution est de 0.205"/pixel.

L'intérêt de VIMOS est son gain multiplex. En mode basse résolution, un spectre s'étale verticalement dans le plan du CCD sur 550 pixels. La longueur des fentes étant de l'ordre de 10 secondes d'arc sur le ciel, soit environ une cinquantaine de pixels, il est possible d'observer simultanément jusqu'à 1000 objets sans que leurs spectres ne se superposent. En mode haute résolution, VIMOS peut imager de l'ordre de 150 spectres simultanément.

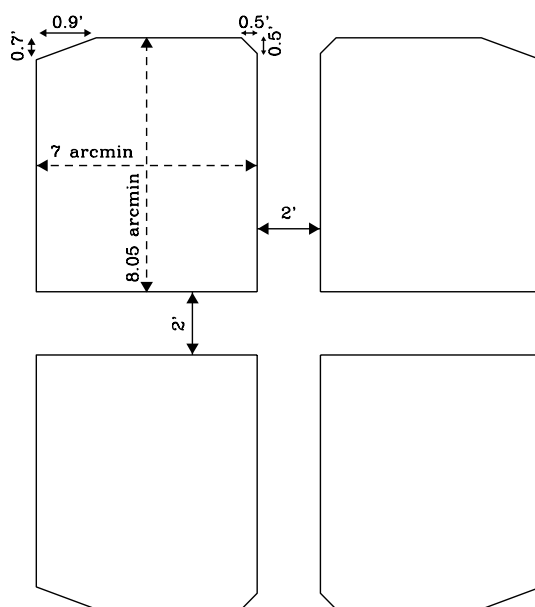


FIG. 2.3 – Champ de vue de l'instrument VIMOS en mode imagerie/spectroscopie multi-objets

2.2.2.2 Spectroscopie intégrale de champ

Le mode spectroscopie intégrale de champ n'a pas été utilisé dans le cadre de cette thèse. On notera cependant que ce mode permet, au moyen de 6400 fibres optiques couplées à des micro-lentilles, d'enregistrer le spectre de chaque zone d'un champ du ciel faisant 57×57 minutes ou 24×24 secondes d'arc carrées. La plage en longueur d'onde disponible est 3700-10000Å. Ce mode permet par exemple l'étude d'amas de galaxies en enregistrant le spectre de tous les objets présents dans le champ.

2.2.3 Observations

Le grand programme VVDS a bénéficié de cinquante nuits d'observations entre octobre 2002 et septembre 2004. Dix-huit ont été perdues suite à des problèmes météorologiques. Dans le cadre de cette thèse, les observations spectroscopiques réalisées avec VIMOS ont été faites en utilisant le grisme rouge basse résolution LR-RED ($R \sim 230$) et le filtre de bande passante 5500-9500Å. La résolution en longueur d'onde du grisme est de 7Å/pixel.

La stratégie d'observation est différente entre la partie *Deep* et la partie *Wide* du sondage, non seulement en terme de magnitude limite de sélection mais aussi en terme d'échantillonnage spatial. Afin de couvrir tous les champs observés, 96 pointés VIMOS ont été définis pour chacun des champs *Wide* VVDS-10h, VVDS-14h et VVDS-22h, 106 dont 66 pointés *Deep* pour le champ VVDS-02h et 5 pour le VVDS-CDFS. Ces pointés se superposent et sont décalés afin de combler le vide créé par la croix séparant les quatre canaux d'un pointé. Les listes des pointés des champs profonds VVDS-02h et CDFS ont été établies afin que chaque zone du ciel soit observée quatre fois, par quatre pointés VIMOS différents (deux fois sur les bords du champ). C'est la stratégie à quatre passages.

Les champs *Wide* ne sont observés qu’au moyen d’une stratégie à deux passages.

La sélection des objets se fait avec le logiciel *Super Spoc* (SSPOC). À une magnitude limite $I_{AB} = 24$, environ un millier d’objets sont visibles dans un quadrant de VIMOS. Le rôle de SSPOC est de maximiser le nombre d’objets qui seront visés simultanément, tout en évitant que leurs spectres ne se chevauchent sur le capteur CCD (voir figures 2.4 et 2.5). Il tient compte pour cela du diamètre apparent de chaque objet et de leur position sur le ciel. Il prend aussi en compte les contraintes liées à la réduction des données (section partie 2.3.1) comme l’ajout d’au moins 1.8 secondes d’arc autour de chaque objet pour enregistrer le spectre du ciel de part et d’autre, et l’alignement des cibles choisies pour permettre un recouvrement des spectres du second ordre.

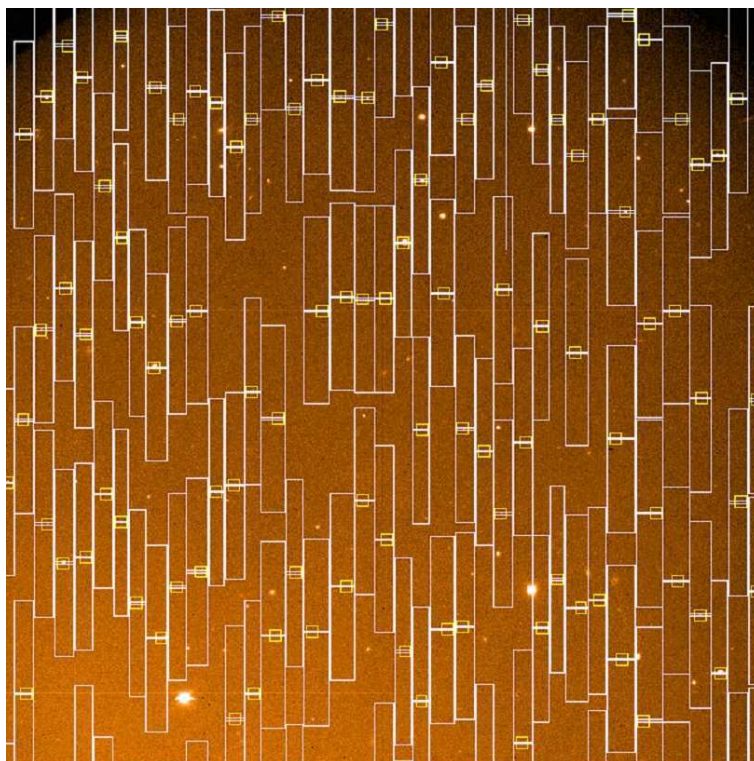


FIG. 2.4 – Exemple d’une sélection par le logiciel *Super Spoc* d’objets pour un quadrant d’une observation MOS. Les cibles sont identifiées par des carrées et positionnées au centre des fentes spectroscopiques alignées selon l’axe Ox . Les zones de dispersion attendues de chaque spectre sont représentées par les grands rectangles alignés verticalement.

Ce programme fait partie d’un ensemble appelé *VLT-VIMOS Mask Preparation Software* (VMMPs, Bottini *et al.* (2005)) destiné à automatiser le processus de sélection des objets et la préparation des observations. Avant chaque observation en mode MOS, une pré-image est obtenue (figure 2.4). Cette dernière, dont le temps d’acquisition est très court (3 minutes), permet de recalibrer les coordonnées des cibles dans le plan focal de VIMOS à partir des objets brillants. Il ne reste plus alors qu’à réaliser le masque constitué des fentes placées devant les différentes cibles grâce au MMU. Plusieurs poses avec VIMOS sont ensuite réalisées pour chaque masque (5 poses de 10 minutes pour les

champs *Wide*, 10 poses de 27 minutes pour la partie *Deep*). Un exemple d'exposition MOS est montrée en figure 2.5 où l'on distingue les spectres enregistrés par VIMOS à travers chaque fente.

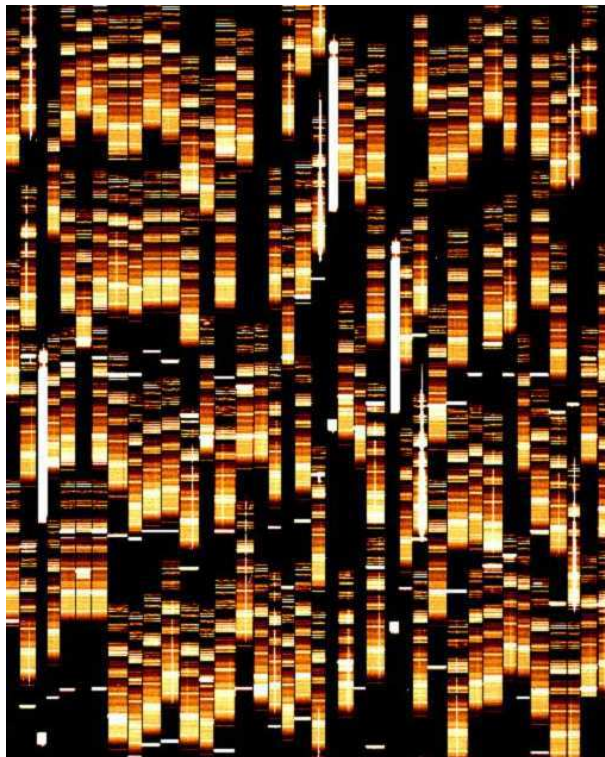


FIG. 2.5 – Exemple d'une exposition MOS typique pour un quadrant de ~ 120 fentes. Les spectres 2D sont alignés selon l'axe Oy

2.3 Réduction des données spectroscopiques et détermination des *redshifts*

Je décris dans cette partie la façon dont les données spectroscopiques obtenues sur le spectrographe multi-objets VIMOS ont été réduites ainsi que la méthode employée pour déterminer le décalage spectral de chaque objet observé. J'ai moi-même participé à la réduction et à la mesure du *redshift* de plus de 6000 objets (5 pointés *Deep* et 9 pointés *Wide*).

2.3.1 Réduction des données

Toute la phase de réduction des données se fait sous le logiciel VIPGI⁴ (Scodeggio *et al.*, 2005) développé à l'IASF de Milan dans le cadre du consortium VVDS afin de réduire la très grande quantité de données (550 spectres par pointé en moyenne pour

⁴voir <http://cosmos.mi.iasf.cnr.it/pandora/vipgi.html>

les observations *Deep*). Il est en effet impossible de réduire autant de données manuellement. Le but de cette étape est de rendre exploitable les données brutes observationnelles et de créer un spectre à une seule dimension représentant le flux émis en fonction de la longueur d'onde. VIPGI est composé de plusieurs modules permettant de réduire les données brutes étape par étape. Des outils de visualisation sont inclus afin de vérifier le bon déroulement de chaque étape et de les valider. Un pointé est constitué de 4 quadrants et la réduction se fait quadrant par quadrant.

De nombreux fichiers sont nécessaires au bon déroulement du processus. Pour chaque masque une dizaine de poses ont été réalisées. À chaque fois, le masque a été décalé de façon à faire glisser les cibles le long de leur fente (axe x). Ceci permet d'améliorer l'étape ultérieure de soustraction du ciel. Des poses de calibration sur des lampes spectrales de référence sont ensuite obtenues pour chaque masque, ainsi que des poses dite *Biais* et *Flat*.

Le *Biais* est la réponse instantanée de la caméra CCD. L'"image" enregistrée correspond au signal à vide du détecteur qu'il faut retirer à chaque pose scientifique. Le *Flat* est l'image enregistrée après une exposition à une lumière à incandescence (corps noir). Cette image permet de corriger la variation de la réponse du CCD de pixel à pixel. Chaque pose scientifique est alors divisée par ce *Flat*, après correction du *Biais*. La pose *Flat* permet aussi ici de localiser précisément la zone de dispersion de chaque spectre sur les images par détection de la zone illuminée.

Chaque pose individuelle d'un spectre est un spectre à deux dimensions : l'axe x est représenté par la fente au milieu de laquelle est placée l'objet (galaxie, étoile, ...) et l'axe y, perpendiculaire, correspond à l'axe selon lequel la lumière qui passe au travers de la fente est dispersée. La figure 2.6 montre un exemple de spectre 2D.

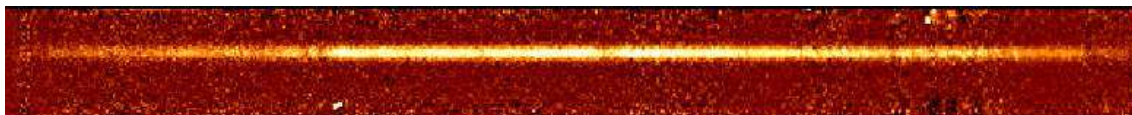


FIG. 2.6 – Spectre à 2 dimensions. L'axe horizontal représente l'axe de dispersion de la lumière, perpendiculairement à l'axe défini par la fente. On observe ici la trace du spectre de la galaxie visée. On identifie sur cet exemple les raies d'absorption CaH et CaK situées au premier quart à gauche du spectre (voir aussi la figure 2.12)

L'addition des poses scientifiques sans tenir compte du décalage suivant l'axe x des fentes permet de réaliser une cartographie des franges d'interférences créées à grande longueur d'onde avec les pixels du détecteur (leur dimension est de 15×15 microns). Ces franges seront retirées par la suite par VIPGI.

Afin d'augmenter le rapport signal à bruit, les spectres sont ensuite additionnés en tenant compte des décalages qui les séparent les uns des autres. En additionnant l'intensité de chacun des pixels le long de l'axe y de dispersion, on crée le profil d'extraction (figure 2.7). Il permet de localiser précisément la position selon l'axe x de la trace du spectre. Les zones de part et d'autre correspondent au spectre du ciel. Celui-ci a été soustrait de chacune des poses individuelles, sur lesquelles se sont superposées les signatures spectrales du ciel (bandes OH), avant co-addition.

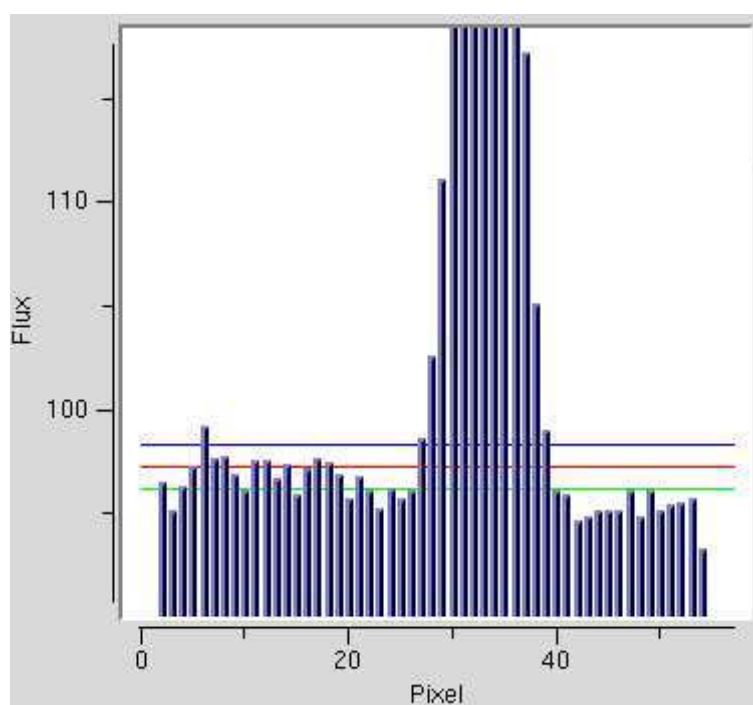


FIG. 2.7 – Profil d'extraction d'un spectre. On a additionné l'intensité des pixels selon l'axe de dispersion du spectre afin de localiser la position exacte du spectre que l'on cherche à extraire.

Entre temps, on aura exploité les spectres obtenus sur des lampes de référence à Hélium-Argon. Les longueurs d'onde de leurs raies en émission sont bien connues. Cela permet au moyen d'une approximation polynomiale de calibrer en longueur d'onde l'axe de dispersion du spectre. Il faut à chaque fois vérifier le fichier de calibration en longueur d'onde avant de l'utiliser. L'écart quadratique moyen entre la position des raies et celle attendue doit être inférieur à 2\AA . S'il s'avère que cette valeur est globalement dépassée, c'est-à-dire pour un trop grand nombre de fentes pour un même quadrant, il faut refaire cette étape avec une autre pose réalisée à partir de lampes spectrales de référence. Il est possible aussi d'utiliser les poses sciences elles-mêmes, mais dans ce cas la tolérance est plus grande (4\AA). VIPGI réalise dans ce cas la calibration à partir des raies du ciel.

On peut alors extraire la position du spectre et obtenir un spectre à une dimension, intensité en fonction de la longueur d'onde. Il reste encore à calibrer en flux en fonction de spectres de référence d'étoiles standards, par exemple LTT3218, puis à appliquer une correction. Cette dernière compense tout d'abord la partie bleue du spectre (autour de 6000\AA) pour laquelle le détecteur est moins sensible. Puis elle corrige des effets d'absorption dues à l'atmosphère terrestre (raie située à $\lambda = 7600\text{\AA}$ par exemple). Ces bandes d'absorption sont clairement visibles sur les spectres d'étoiles de type K dont le spectre est assez plat.

Enfin, la dernière étape consiste à nettoyer les spectres. La réduction n'est en effet pas toujours parfaite à cause d'une mauvaise correction des raies d'interférences dans la partie rouge du spectre. Ces aléas introduisent des artefacts dans le spectre final pouvant être interprétés à tort comme des raies d'émissions ou d'absorptions. Ceci peut alors fausser la mesure du *redshift* (partie 2.3.2.2). De plus, il arrive que deux spectres 2D soient

parfaitement alignés sur le CCD dans le plan focal de VIMOS. Si l'un de ces spectres correspond à un objet brillant (étoile de magnitude $I_{AB} = 18$ par exemple), l'ordre zéro de son spectre (son image) peut se retrouver superposé au deuxième spectre. Il n'est malheureusement pas possible de supprimer cet ordre de façon automatique lors des étapes précédentes de réduction. Il faut donc le supprimer à la main, spectre par spectre lorsque la situation l'exige.

Pour terminer, lors de la prise de données avec VIMOS, il arrive qu'un deuxième objet (exceptionnellement trois), non sélectionné au préalable, soit vu à travers la fente de l'objet ciblé. Deux spectres sont alors enregistrés sur le CCD. VIPGI est capable d'extraire les deux spectres dès que leur intensité les rends détectables. Le deuxième objet peut ou non vérifier les critères de sélection de l'échantillon complet. Dans tous les cas il faut l'identifier dans la base de donnée VVDS et il sera répertorié de façon à garder la trace qu'il ne s'agissait pas de l'objet visé (voir partie 2.3.2.3). Dans certain cas, l'objet n'est pas réel. Il s'agit généralement du second ordre ou de l'ordre zéro d'un autre spectre.

2.3.2 Détermination des *redshifts*

Après avoir réduit les données et obtenus les spectres à une dimension calibrés en longueur d'onde, en flux, et nettoyés l'étape suivante consiste à mesurer le décalage spectral de chacun d'entre eux. Le décalage vers le rouge, ou *redshift*, est défini par $z = \lambda/\lambda_0 - 1$, avec λ la longueur d'onde à laquelle est reçue le flux émis à la longueur λ_0 . Dans le cadre du VVDS, les objets sont observés dans l'intervalle [5500–9500Å]. Il faut donc identifier les signatures spectrales observées dans cet intervalle pour mesurer la valeur de z . Ceci est d'autant plus difficile que l'on a aucune connaissance à priori du *redshift*, aucune préselection en couleur et/ou morphologie n'ayant été faite. Des étoiles ainsi que des quasars (ou QSO pour *quasi stellar object*) sont donc présents dans l'échantillon. De même, de part la profondeur de l'échantillon ($I_{AB} = 24$), de nombreuses galaxies à grands décalages spectraux (jusqu'à $z \sim 5$) sont présentes.

2.3.2.1 Raies en émission, en absorption

Les galaxies ciblées dans le VVDS sont des galaxies lointaines pour la plupart, donc de dimensions angulaires très petites (de l'ordre de 1 à 2 secondes d'arc), et parfois dominées par le *seeing*. Ainsi les spectres mesurés sont des spectres intégrés de toutes les étoiles contenues dans ces galaxies. Selon l'évolution de ces dernières, le spectre total sera donc différent. On distingue trois types de spectres :

spectres de raies d'émission

Les raies d'émission correspondent à des pics d'intensité dans le spectre. Elles peuvent être plus ou moins intenses et plus ou moins larges. Elles se produisent lors du passage d'un électron d'un niveau d'énergie atomique E_2 à un autre moins énergétique E_1 libérant un photon de longueur d'onde λ définis par :

$$\lambda = \frac{hc}{E_2 - E_1} \quad (2.1)$$

où h est la constante de Planck⁵ et c la vitesse de la lumière. Cette énergie peut aussi être dissipée par collisions si la densité est suffisante, ce qui n'est généralement pas le cas dans le milieu interstellaire des galaxies. On observe ainsi de nombreuses raies dites *interdites* ([OII], [OIII], ...) invisibles sur Terre où la densité est $\sim 10^{16}$ fois plus importante que les conditions requises.

spectres de raies d'absorption

La présence d'une raie d'absorption indique qu'il y a entre la source lumineuse et l'observateur un milieu absorbant à cette longueur d'onde. Comme pour les raies d'émission, les raies d'absorption peuvent être plus ou moins intenses et plus ou moins larges. Elles se produisent lors du passage d'un électron d'un niveau d'énergie atomique à un autre plus énergétique, grâce à l'absorption d'un photon.

spectres continus

Un spectre continu est observé lorsqu'une source émet dans un large domaine de longueur d'onde, de manière continue. C'est le cas du spectre de corps noir émis par l'atmosphère des étoiles. Il présente un maximum duquel on déduit la température de l'étoile. Cette température se traduit par une couleur et une durée de vie de l'étoile. Les étoiles bleues ayant une durée de vie très courte ($\sim 10^6$ années), la présence d'un maximum dans le bleu dans le spectre continu d'une galaxie est donc caractéristique des galaxies où une phase de formation stellaire récente a eu lieu.

Le tableau 2.3 rapporte une liste non-exhaustive de raies spectrales. Il y est indiqué si ces raies peuvent être observées en émission (E), ou en absorption (A). Certaines raies peuvent être très larges (B). Elles sont produites par des jets violents de matière dans la direction de la ligne de visée. Ces derniers sont généralement produits par un trou noir présent dans la galaxie hôte et caractéristique des quasars. La figure 2.8 indique l'intervalle en *redshift* d'observabilité de quelques une de ces raies dans l'intervalle en longueur d'onde [5500-9500Å].

De nombreuses références de spectres de galaxies existent. On notera les quatre spectres de galaxies observées dans l'Univers local par Coleman *et al.* (1980) (figure 1.10), ou encore les spectres rapportés par Kennicutt (1992). Il est difficile d'obtenir des spectres de référence émis dans l'ultraviolet par des galaxies locales. Par contre cela devient possible si l'on observe à grand *redshift*. Ainsi, Steidel *et al.* (1996) et Shapley *et al.* (2003) rapportent des exemples de spectres de galaxies observées à $z \sim 3-4$. À de tels décalages vers le rouge, le flux lumineux émis par une galaxie en UV est reçu dans le visible. Une analyse spectroscopique est alors possible.

Des spectres de référence ont été créés à partir des spectres du VVDS et inclus dans l'outil VIPGI (voir section 2.3.2.2). L'ensemble des spectres VVDS a été rendu disponible au consortium en fin de cette thèse. La figure 2.9 montre 6 de ces spectres de références (*templates*).

Quelques exemples de spectres composites de quasars ont été obtenus à partir de données du *Hubble Space Telescope* (HST) par Zheng *et al.* (1997), ou plus récemment, dans le cadre du *Sloan Digital Sky Survey* (SDSS) par Vanden Berk *et al.* (2001) et Budavári *et al.* (2001).

⁵ $h = 6.63 \times 10^{-34}$ J.s

λ (Å)	type	nom	λ (Å)	type	nom	λ (Å)	type	nom
1215.7	A,E,B	Ly α	2626.0	E	FeII	4304.4	A	bande G
1260.4	A	SiII	2799.0	A,E,B	MgII	4340.4	A,E	H γ
1303.0	A	[OI]	2852.9	A	MgI	4363.2	E	[OIII]
1334.5	A	CII	3345.9	E	[NeV]	4383.6	A	FeI
1397.0	A,E,B	SiIV	3425.8	E	[NeV]	4861.3	A,E,B	H β
1526.7	A	SiII	3581.0	A	FeI	4958.9	E	[OIII]
1549.0	A,E,B	CIV	3727.5	E	[OII]	5006.8	E	[OIII]
1608.5	A	FeII	3797.9	A,E	H θ	5175.4	A	MgI
1640.0	A,E	HeII	3868.7	E	[NeIII]	5269.0	A	Ca+Fe
1856.0	A	AlIII	3883.0	A	CN	5875.6	A,E	HeI
1909.0	E,B	[CIII]	3933.7	A	CaII K	5892.5	A	Na-bl
2344.2	A	FeII	3968.5	A	CaII H	6562.8	A,E,B	H α
2382.0	A	FeII	3967.5	E	[NeIII]	6583.6	E	[NII]
2586.7	A	FeII	4101.7	A,E	H δ	6678.1	A,E	HeI
2600.0	A	FeII	4216.0	A	CN	6725.0	E	[SII]

TAB. 2.3 – Liste des principales raies en émission (E), et/ou absorption (A) en fonction de leur longueur d'onde au repos. Les raies entre crochets sont dites *interdites*. Les raies pouvant être larges sont notées B.

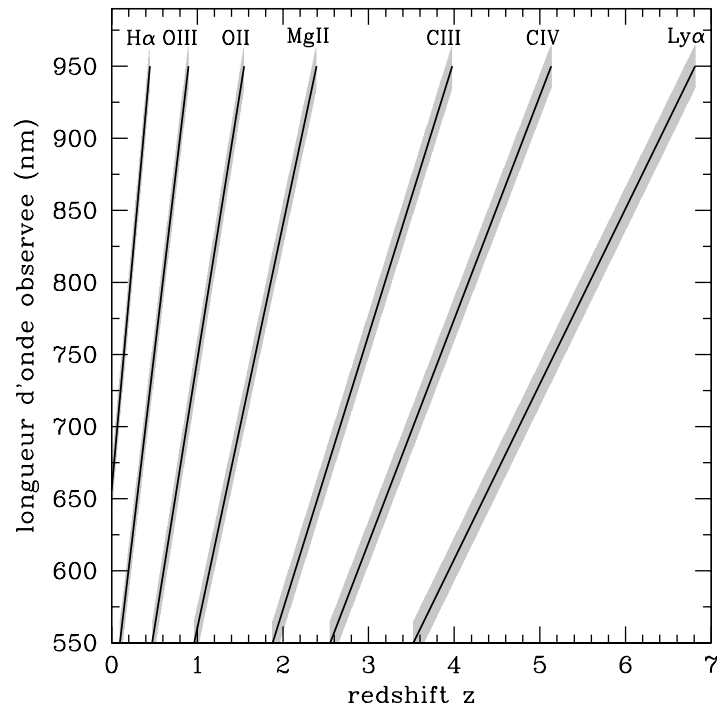


FIG. 2.8 – Visibilité de quelques raies en fonction du *redshift*. Par exemple, la raie MgII n'est observable qu'entre $z=0.97$ et 2.39 . De même, les raies [OII] et [OIII] peuvent être simultanément observées sur le spectre d'une galaxies à $z=0.6$

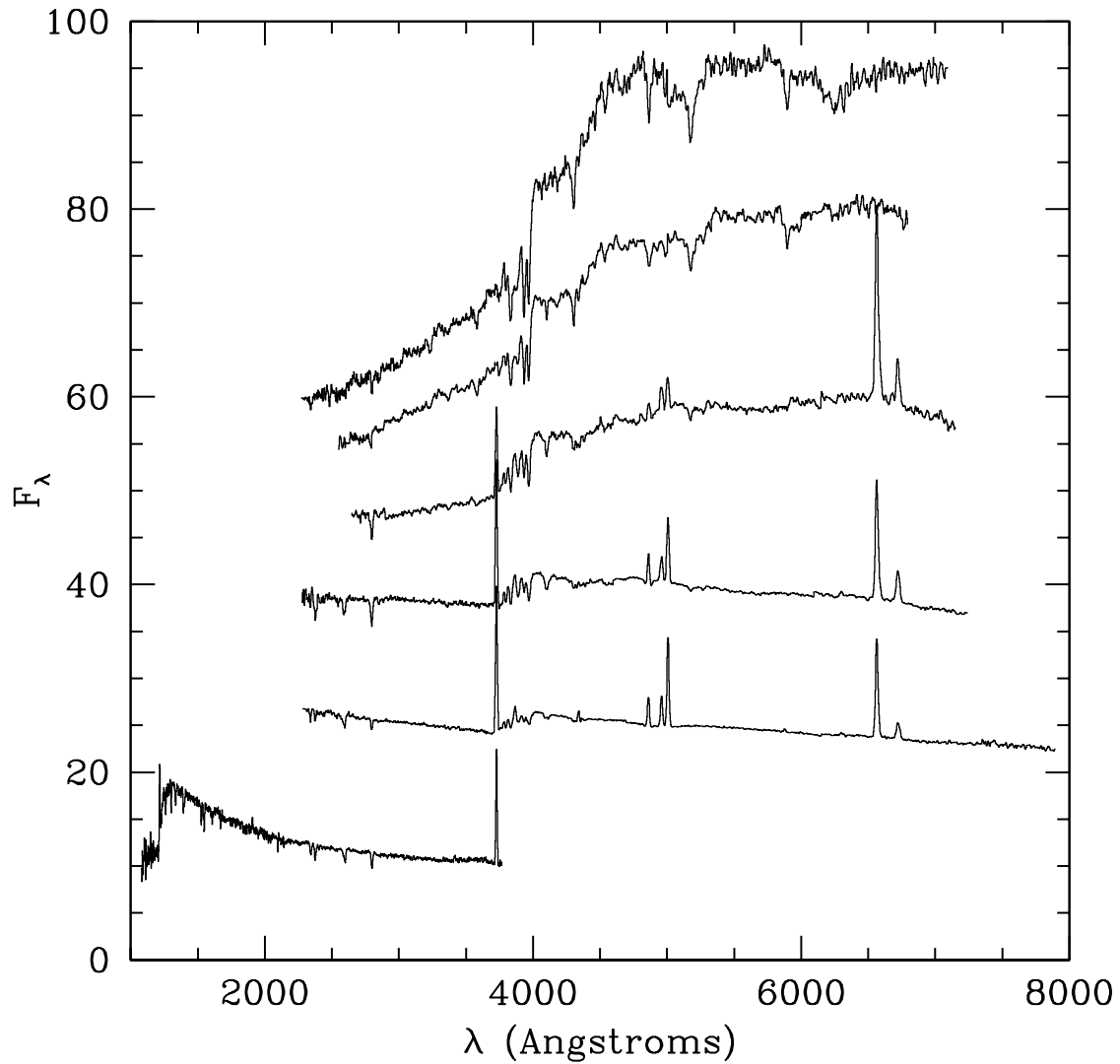


FIG. 2.9 – Spectres de référence réalisés à partir des spectres du VVDS. On retrouve un *template* composé à partir des spectres de galaxies les plus rouges, puis un template pour chacun des type 1, 2, 3 et 4 (voir définition des types partie 5.2.1) et enfin un *template* composé des spectres de galaxies dont le *redshift* est compris entre 1.4 et 5. Les 5 premiers ont été réalisés à partir de spectres dont la note de qualité du *redshift* (partie 2.3.2.3) est de 3 ou 4. Les notes 2 ont aussi été utilisées pour la réalisation du sixième.

Enfin, une référence complète de spectres d'étoiles est proposée par [Pickles \(1985\)](#).

2.3.2.2 Méthodologie

L'étape cruciale de mesure des *redshifts* est loin d'être évidente. Aucun entraînement n'ayant pu être possible sur des spectres simulés, l'apprentissage de la mesure des décalages spectraux s'est fait de manière itérative pour la plupart des membres du VVDS à partir des données réelles obtenues avec VIMOS. De plus, de part la très grande quantité de données, un processus "semi-automatique" a dû être mis en place.

Pour cela, plusieurs outils ont été développés pour faciliter l'identification des raies spectrales, la décision finale de la valeur d'un *redshift* restant entre des mains humaines. De plus, pour s'assurer d'une homogénéité des résultats d'un jeu de données à un autre, les *redshifts* de chaque quadrant ont été mesurés indépendamment par deux personnes différentes. Les deux listes de *redshifts* ont ensuite été comparées par ces deux mêmes personnes. Une décision finale est ensuite prise dans les cas de désaccord.

La démarche suivie par chaque personne impliquée dans la mesure des *redshifts* est la suivante.

Premier passage : mesure automatique

Le programme KBRED développé par [Scaramella et l'équipe VVDS \(2005\)](#) utilise des techniques de corrélation entre les spectres VIMOS et une bibliothèque de spectres de références pour déterminer leurs *redshifts*. Cet outil produit sous IDL est accessible directement à l'utilisateur via l'interface de VIPGI. Il peut être utilisé au cas par cas ou bien sur tous les spectres d'un quadrant VIMOS.

À partir d'une librairie de spectres de références (*templates*) au repos, KBRED génère toute une série de spectres à différents *redshifts* et recherche la meilleure solution possible. Le critère d'arrêt correspond au minimum d'un test du χ^2 , résultat de la corrélation croisée entre le spectre VIMOS et les *templates*.

Ce processus est donc dans un premier temps lancé en aveugle sur tous les spectres réduits d'un quadrant VIMOS. Il en ressort une première estimation du décalage vers le rouge des spectres. Les étoiles identifiées comme telles par KBRED ont leur *redshift* mis à zéro. Les résultats sont écrits dans un fichier directement utilisable par VIPGI par la suite.

Malheureusement, KBRED n'a pas un taux de succès de 100%. À posteriori, seulement ~60% des *redshifts* sont bien mesurés par cette approche automatique. Quoi qu'il arrive, KBRED donnera toujours une solution. Certaines des mauvaises solutions s'expliquent donc par

1. un mauvais nettoyage du spectre 1D : KBRED interprète alors des ordres zéro ou des résidus de soustraction des raies du ciel comme des signatures spectrales.
2. l'absence de signatures noyées soit dans le bruit, soit trop faiblement marquées
3. un manque de *templates* correspondant à l'objet observé

Il est donc nécessaire de regarder une par une toutes les solutions de KBRED pour corriger les erreurs les plus grossières et travailler de façon itérative sur les spectres les plus délicats (spectres très bruités, sans raie d'émission/absorption clairement identifiable).

Deuxième passage : les spectres un par un

L'interface graphique de VIPGI (figure 2.12) permet à l'utilisateur d'afficher la position des raies attendues pour un *redshift* z donné. Cette facilité permet en un coup d'oeil de tester la crédibilité d'une valeur de z en comparant le spectre 2D réduit, le spectre 1D et le spectre du ciel. Si la solution proposée par KBRED à l'étape précédente n'est pas convaincante pour l'une des raisons suivantes :

- étoile K confondue avec un spectre de galaxie elliptique à $z \sim 0.4$ (ou vice-versa),
- couple évident de raies clairement visibles mais oubliées par KBRED,
- continu du spectre en désaccord avec la solution proposée
- solution basée sur une signature fausse (ordre zéro, raie du ciel, . . .)

alors l'utilisateur dispose de plusieurs outils pour essayer de déterminer le bon *redshift*.

Si une ou plusieurs raies (émission et/ou absorption) apparaissent aux yeux de l'utilisateur, il peut indiquer leurs positions à VIPGI. En utilisant la liste de raies incluses dans VIPGI, il restera à tester les solutions possibles jusqu'à la découverte de la bonne solution.

Si aucune signature spectrale claire n'est présente (à cause d'un rapport signal à bruit trop faible par exemple), mais que le spectre a visiblement été correctement extrait, il est possible de réutiliser KBRED. Dans un mode manuel, il est possible de choisir la librairie de *templates*, la plage en longueur d'onde à utiliser pour la corrélation, ou encore l'intervalle en *redshift* où chercher la meilleure solution. Cette façon de travailler permet de détailler les meilleures solutions possibles et de converger vers la meilleure solution. Les figures 2.10 et 2.11 montrent deux exemples de solutions proposées par KBRED. Chaque panneau indique la meilleure solution trouvée par KBRED dans les différents intervalles en *redshift* que l'utilisateur veut tester. Parmi ces différentes solutions, KBRED indique celle qui est la plus crédible (*BEST P²*).

Dans tous les cas, il est nécessaire de donner une note à chaque *redshift* afin d'être capable a posteriori d'écarter d'un sous échantillon de galaxies celles pour lesquelles le *redshift* est peu sur. Cette notation est détaillée dans la partie 2.3.2.3.

2.3.2.3 Indices de confiance

Chaque *redshift* mesuré se voit attribué une note. Celle-ci traduit la confiance qu'il faut accorder à cette mesure. Elle peut évoluer lorsque les deux mesures réalisées indépendamment sont confrontées. Cette note, introduite par [Le Fèvre et al. \(1995\)](#) dans le cadre du sondage spectroscopique CFRS, rassemble en fait beaucoup plus d'information que la seule probabilité que la mesure du *redshift* soit correct. Elle permet en effet d'identifier dans l'échantillon complet les spectres de quasars, les seconds objets non visés initialement mais identifiés dans la base de donnée, ceux non identifiés, les spectres inexploitable pour diverses raisons, les fausses détections. Les définitions de ces notes, lorsque l'objet est présent dans la base de donnée, sont rassemblées dans le tableau 2.4. Les notes des objets non identifiés se voient rajouter le signe — devant.

Dans certains cas, aucun *redshift* ne peut être mesuré. Lorsque le spectre à deux dimensions montre clairement la présence de l'objet dans la fente, mais que la mesure est impossible (signal très bruité ou trop faible, présence d'aucune signature spectrale sur laquelle baser un *redshift*) la note attribuée est 0 (20 dans le cas d'un second objet). Lorsqu'il s'agit d'un problème technique (mauvaise réduction, objet pas observé), la note est alors de -10. S'il s'agit d'une fausse détection (ordre zéro, second ordre, spectre du

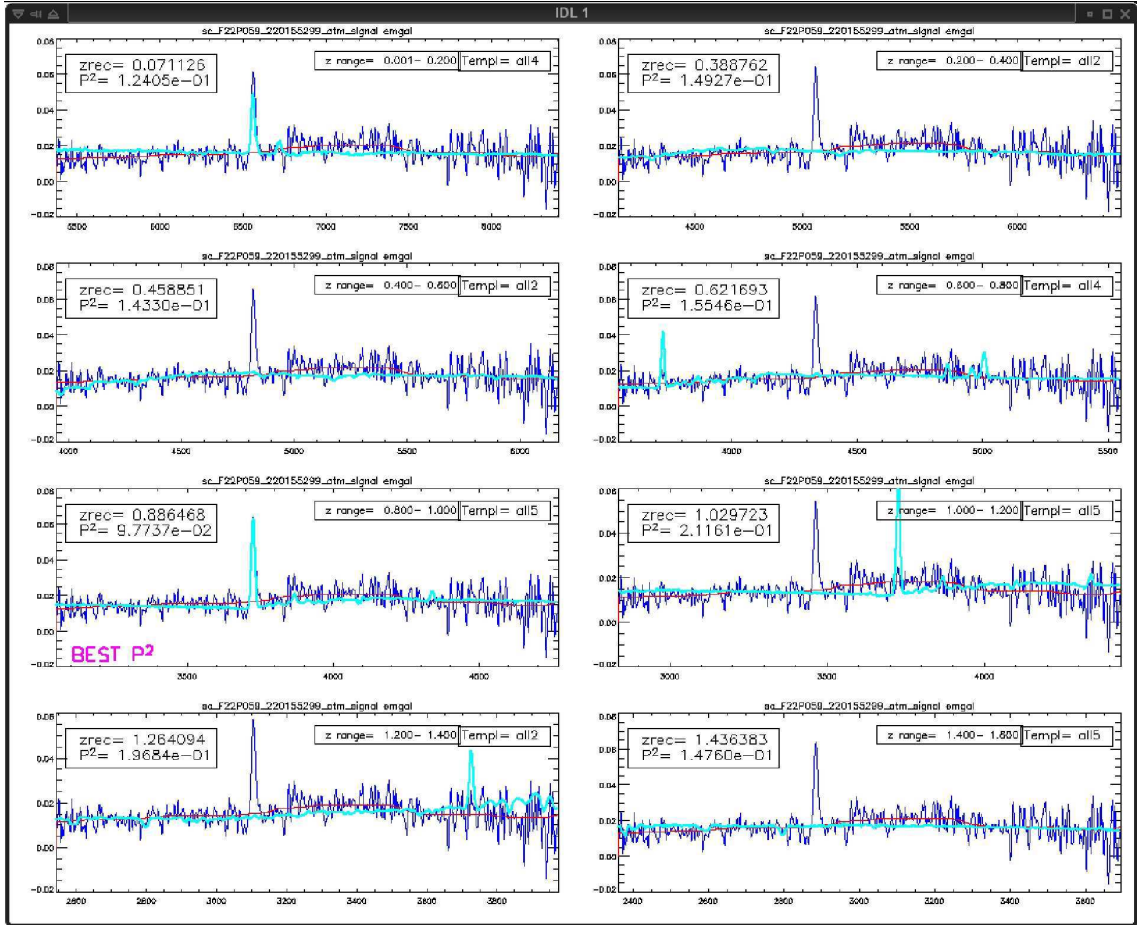


FIG. 2.10 – Solutions proposées par KBRED dans les 8 intervalles en *redshift* définis par l'utilisateur. La meilleure solution identifiée par KBRED est celle marquée *BEST P²*. Dans cet exemple, KBRED a identifié une galaxie à $z=0.8865$. On distingue la raie en émission [OII].

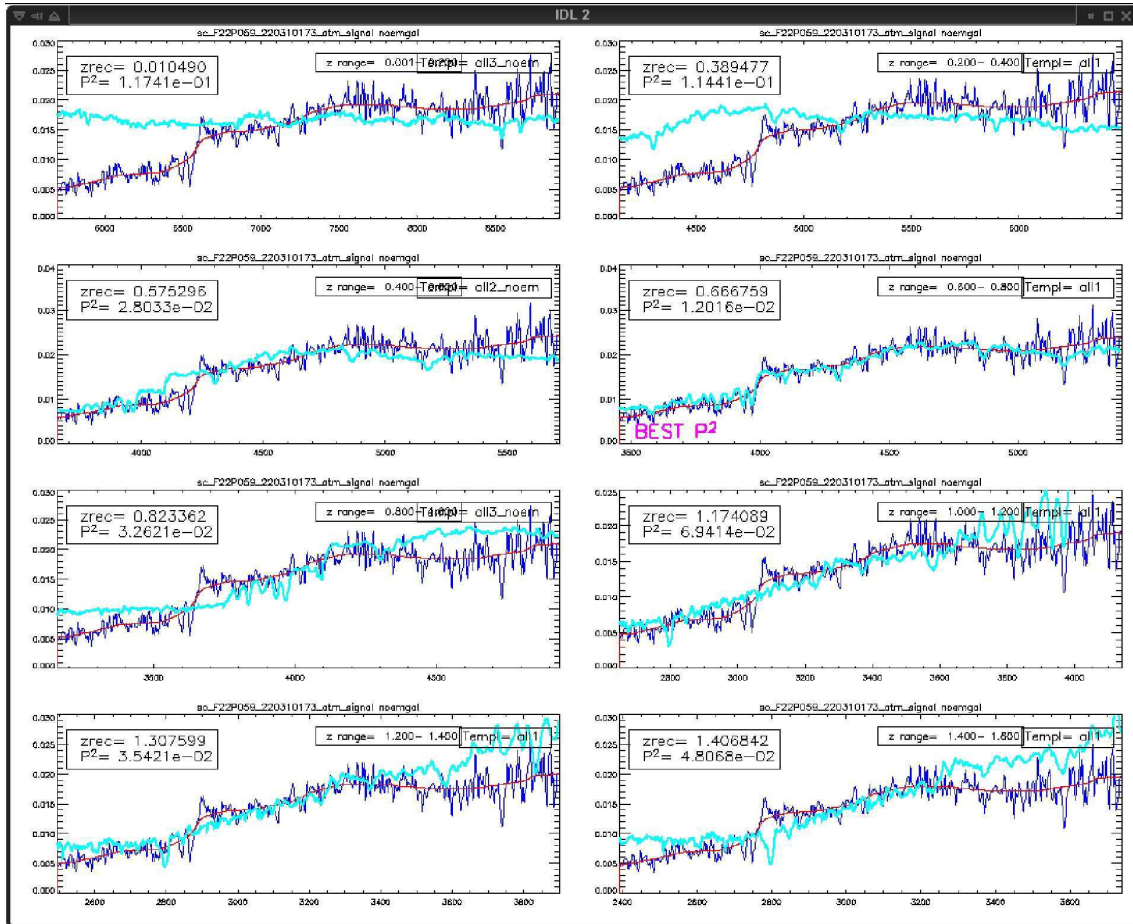


FIG. 2.11 – Solutions proposées par KBRED dans les 8 intervalles en *redshift* définis par l'utilisateur. La meilleure solution identifiée par KBRED est celle marquée *BEST P²*. Dans cet exemple, KBRED a identifié une galaxie à $z=0.6668$. On distingue par exemple les raies en absorption CaH et CaK

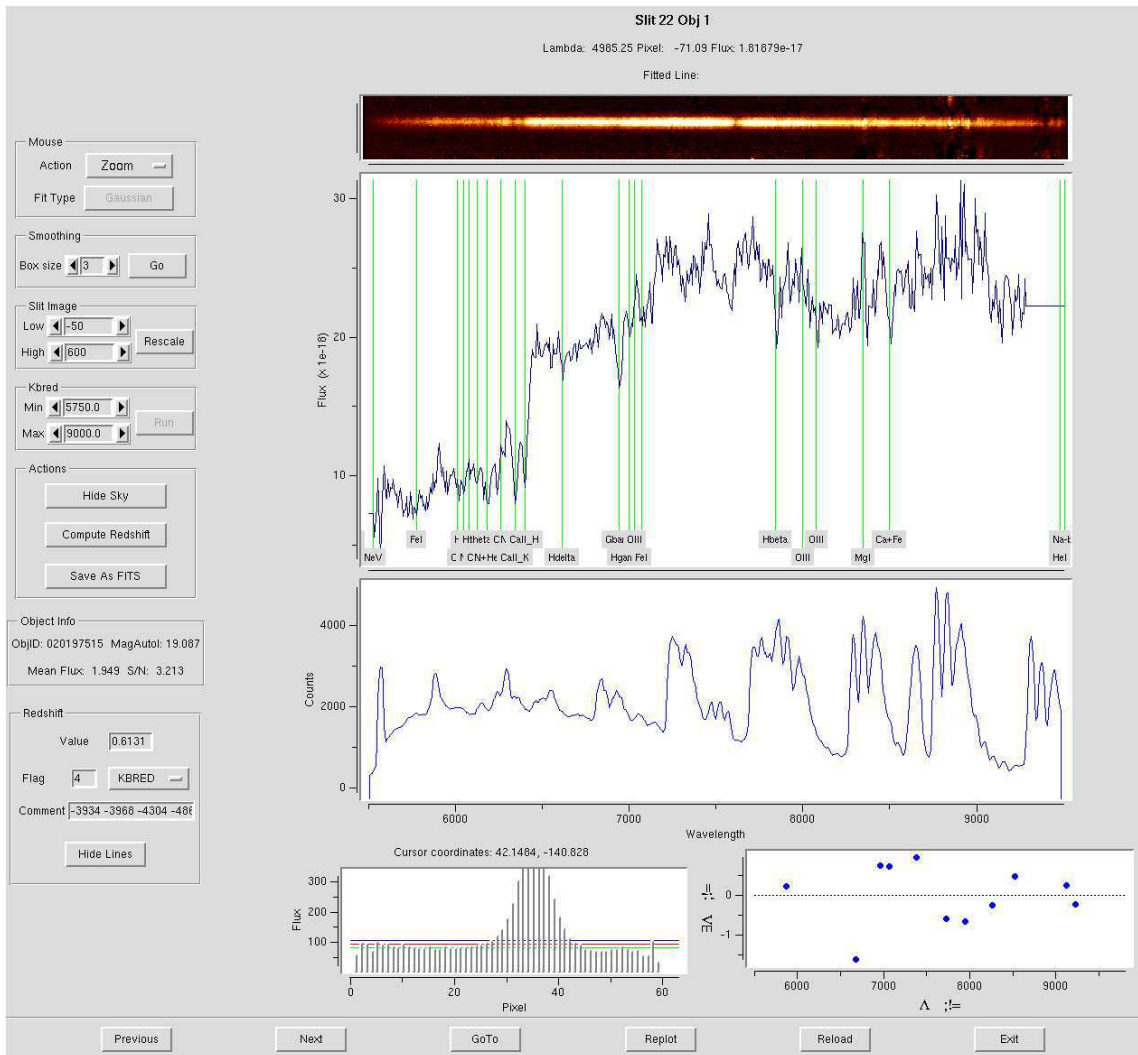


FIG. 2.12 – Panneau *Slit Summary* de VIPGI. Permet une visualisation instantanée du spectre 2D réduit (en haut), du spectre 1D (au milieu), et ici du spectre du ciel (en bas en bleu). Cette interface permet une vérification et/ou mesure du *redshift* indiqué en bas à gauche du panneau, où l'on retrouve également sa note de qualité (voir partie 2.3.2.3) ainsi que le listing des raies identifiées.

TAB. 2.4 – Définition des notes de confiance attribuées à chaque *redshift* dont l'objet est identifié dans la base de donnée.

degré de confiance dans le <i>redshift</i>	objet visé		second objet		commentaires
	galaxies étoiles	quasars	galaxies étoiles	quasars	
50%	1	11	21	211	faible RSB ^a et/ou peu de raies identifiées
75%	2	12	22	212	RSB moyen, et/ou peu de raies identifiées
95%	3	13	23	213	plusieurs raies, forme du continu
100%	4	14	24	214	pas de doute possible
une seule raie	9	19	29	219	la forme du continu ne permet pas de conclure

^aRapport Signal à Bruit

ciel mal soustrait en bord de fente), la note est -99. Dans tous ces cas la valeur du *redshift* est mise par convention à $z=9.9999$. Cette notation permet ensuite de traiter facilement l'ensemble de l'échantillon en appliquant des choix clairs qui conduisent à une analyse statistique robuste.

2.3.2.4 Exemple de spectres

Les figures 2.13 et 2.14 montrent des exemples de spectres du champ VVDS-02h. Sur la première, on distingue, de haut en bas et de gauche à droite

- un spectre de galaxie de type spirale à bas *redshift* présentant des raies en émission de H β , [OIII], H α et [SII]
- un spectre de galaxie de type elliptique avec un fort *break* à 4000Å observé à $\sim 7100\text{\AA}$ précédé des raies d'absorption H et K du calcium
- un spectre avec une forte raie [OII] en émission
- un spectre à $z=1.7225$ identifié par les raies en absorption du fer FeII et du magnésium MgII
- un spectre à $z=3.3001$ identifié sur la base de la raie CIV et autres raies UV en absorption
- un spectre de galaxie à très grand décalage spectral ($z=3.8725$) pour lequel on observe la raie Lyman α en émission

La figure 2.14 montre des exemples de spectres d'étoiles (de type K en haut à gauche et de type M en haut à droite) et des exemples de spectres de quasars présentant des raies en émission très larges (Lyman α en bas à gauche, et MgII en bas à droite).

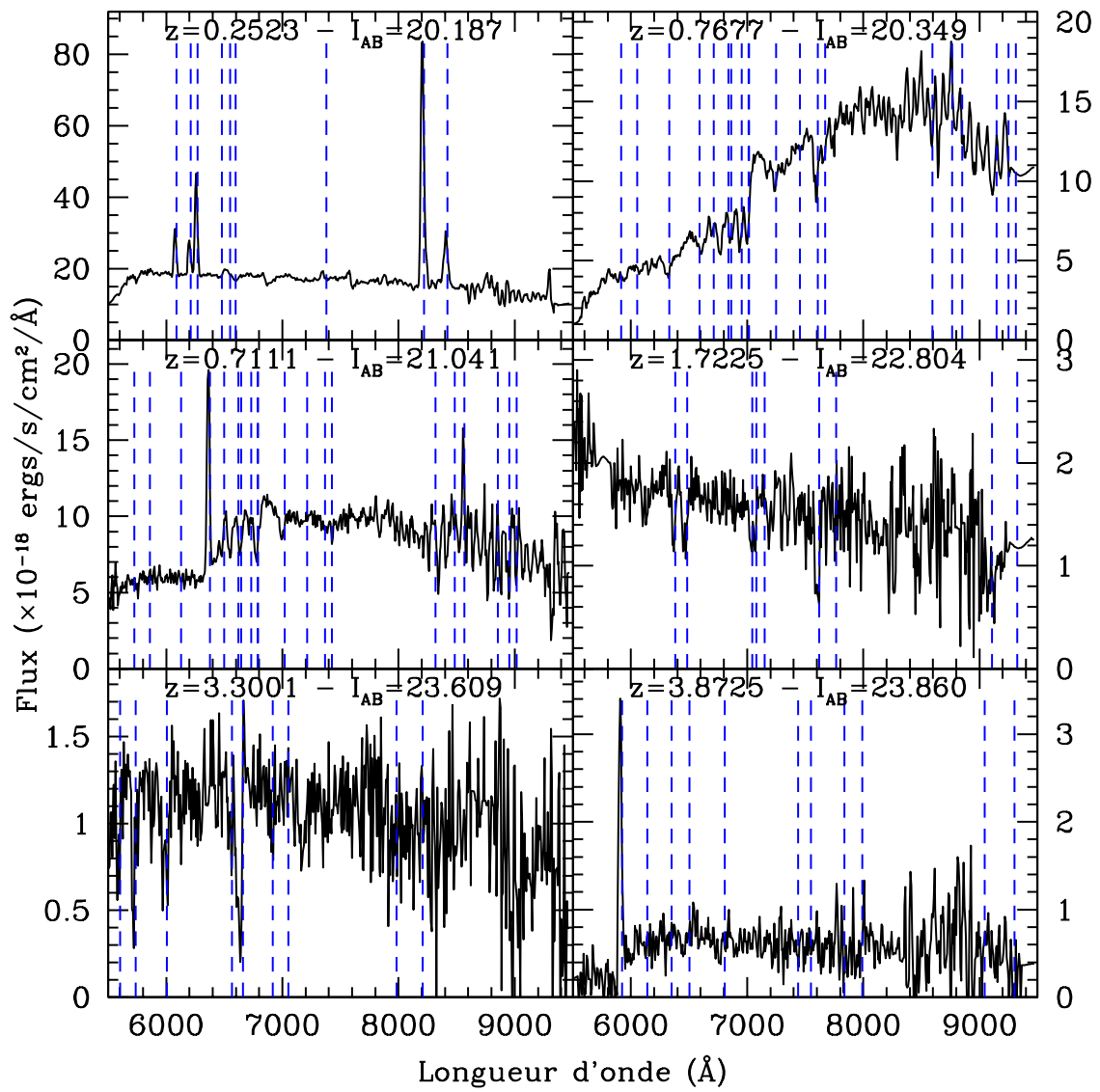


FIG. 2.13 – Spectres de galaxies du VVDS-02h

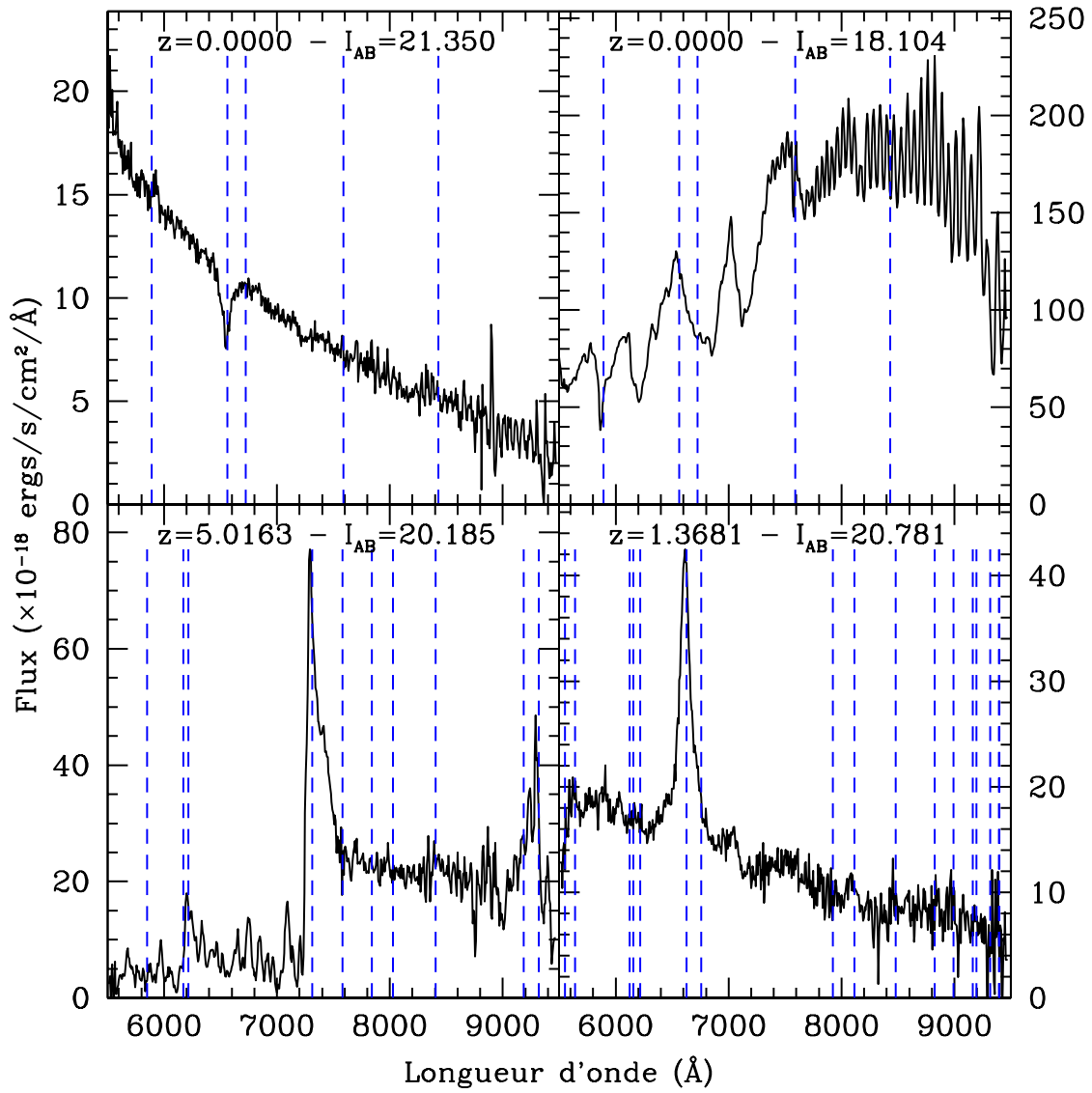


FIG. 2.14 – Spectres d'étoiles et de quasars du VVDS-02h. On observe les résidus de correction de *fringing* sur les objets très brillants dans la partie rouge du spectre (en haut à droite)

2.4 Première campagne d'observation : les données

Lors de la première campagne d'observation, dans le cadre des nuits garanties au consortium, la priorité a été donnée à l'observation des champs profonds VVDS-02h et CDFS. Quelques observations ont tout de même été réalisées sur les champs larges VVDS-10h et VVDS-22h.

2.4.1 Les champs *Deep*

Les objets observés dans les champs *Deep* sont généralement peu lumineux. Leurs magnitudes vont en effet jusqu'à $I_{AB} = 24$. Les temps de poses sont donc long : 10 poses de 27 minutes par pointé soit 4h30 d'intégration. Cette partie décrit les données *Deep*. On se reportera aux articles de [Le Fèvre et al. \(2004b\)](#) et [Le Fèvre et al. \(2005c\)](#) pour plus d'informations.

2.4.1.1 Le VVDS-02h

La surface du champ VVDS-02h couverte lors de la première époque d'observation est de 0.82×0.68 degrés carrés, ce qui est sans précédent à cette profondeur $I_{AB} = 24$. Elle a été couverte par 20 pointés VIMOS. Chaque partie de la région centrale a été observée par 4 pointés différents afin d'augmenter l'échantillonnage spatial. La figure 2.15 représente les objets spectroscopiques du champ VVDS-02h. Les quadrants des 20 pointés VIMOS y sont représentés en rouge. Les zones vides sont celles éliminées du catalogue photométrique (étoiles saturées) avant sélection des objets. On observe clairement la différence de densité entre la région centrale à 4 passages et celle extérieure à seulement 2 passages. Le taux d'échantillonnage est ainsi de $\sim 40\%$ de l'ensemble des sources photométriques détectées dans cette zone. Il est en 22.8% sur l'ensemble de la zone couverte.

La distribution en *redshift* de l'échantillon à $I_{AB} = 24$ est représenté en figure 2.16. Le *redshift* moyen est $\langle z \rangle \sim 0.9$ avec une queue de distribution qui s'étire jusqu'à $z \sim 5$. On compte 419 *redshifts* au-delà de $z=2.5$ pour des notes de confiance comprises entre 1 et 9 (169 avec uniquement les notes de 2 à 9). On observe un intervalle en *redshift*, $2.2 \leq z \leq 2.7$, qui contient très peu d'objets. Cet intervalle est le "*Désert en Redshift*" ([Le Fèvre et al., 2005c](#)). Les spectres des objets situés à ces décalages spectraux présentent peu de signatures spectrales observables et identifiable entre 5500 et 9500Å. La présence des fortes raies OH du ciel aux alentours de 9000Å complique la détermination de ces *redshifts*. On peut alors estimer qu'une large fraction de ces objets se voient attribuer une note de 0 ou un mauvais *redshift*. Ils n'apparaissent donc pas dans la distribution en *redshift*. Cette distribution est utilisée dans le calcul de la fonction de corrélation afin de répartir sur l'axe du *redshift* les objets du catalogue de points générés aléatoirement..

L'erreur commise sur la mesure des *redshifts* a été évaluée à partir de 426 objets dont les spectres ont été observés deux fois. Ainsi la comparaison de ces *redshifts* indique une précision de 276 km.s^{-1} .

La table 2.5 répartit les 9842 objets du champ profond VVDS-02h pour lesquels un spectre a été enregistré avec VIMOS en fonction de leur nature (étoiles, galaxies, quasars) et de la note de confiance attribuée au *redshift* mesuré.

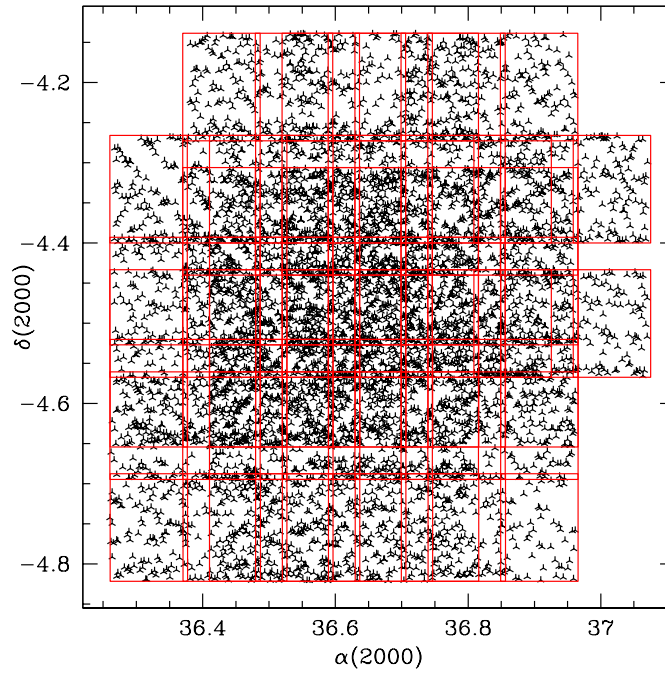


FIG. 2.15 – Objets observés avec VIMOS dans le VVDS-02h. Les quadrants des 20 pointés VIMOS sont représentés en rouge. Un objet de la région centrale a eu quatre chances d’être sélectionné (4 passages), deux chances pour les parties extérieures.

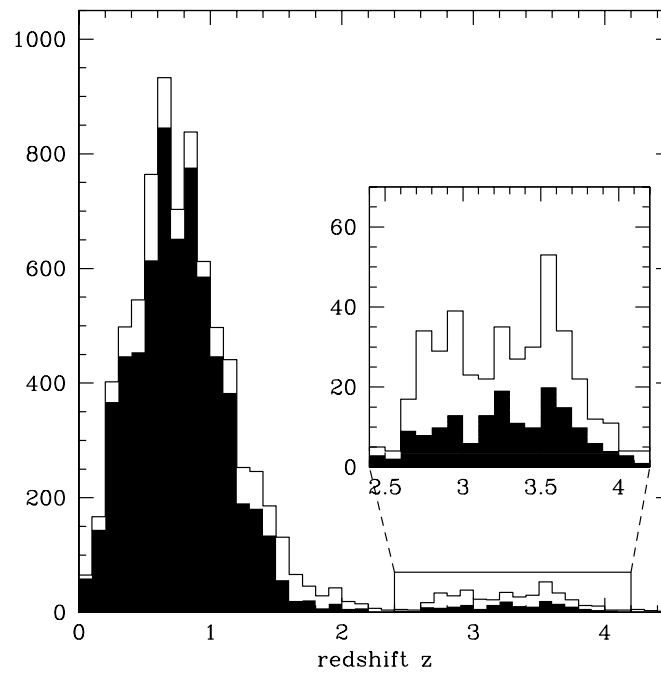


FIG. 2.16 – Distribution en *redshift* du VVDS-02h ($17.5 \leq I_{AB} \leq 24$). En noir la distribution des notes 2 à 9, en blanc celle en ajoutant les notes 1.

	galaxies		étoiles		quasars		non identifié
note	1	2-9	1	2-4	11	12-19	0
nombre	1363	6582	63	623	13	62	690
note	21	22-29	21	22-24	211	212-219	20
nombre	72	209	1	10	0	5	149

TAB. 2.5 – Répartition du nombre d'objets en fonction de leur nature et de leur note dans le VVDS-02h

On peut définir le taux de succès spectroscopique comme le rapport entre le nombre de *redshifts* mesurés et le nombre d'objets observés. Comme attendu, ce taux dépend de la magnitude des objets. La figure 2.17 montre en effet qu'il est plus difficile de mesurer le *redshift* d'un objet faible que d'un objet brillant. Si l'on restreint l'échantillon aux notes 2 à 9, il n'a été possible d'identifier le *redshift* de seulement 55% des objets de magnitude I_{AB} comprise entre 23.5 et 24. Par contre ce taux monte à près de 85% si l'on inclus les notes 1.

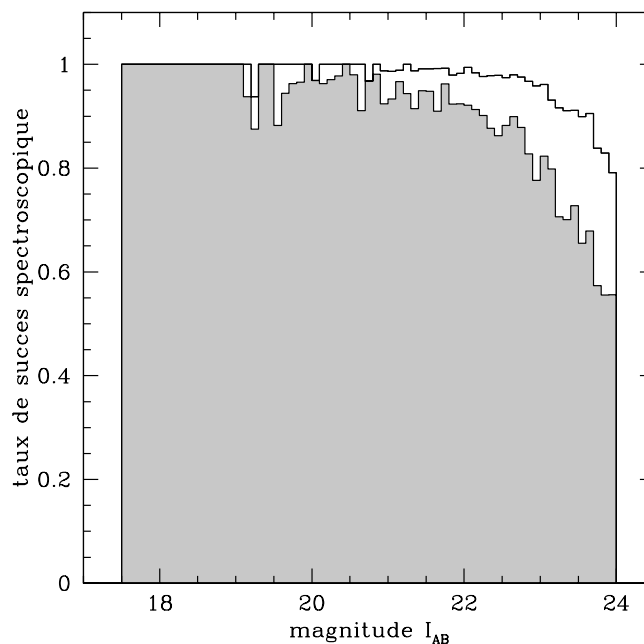


FIG. 2.17 – Le taux de succès spectroscopique dépend de la magnitude des objets. L'histogramme en gris correspond au rapport entre le nombre de *redshifts* de note 2 à 9 et le nombre de *redshifts* de note 0 à 9, en fonction de la magnitude. L'histogramme blanc tient compte des notes 1.

2.4.1.2 Le *Chandra Deep Field South*

Les données prises dans le *Chandra Deep Field South* (CDFS, (Giacconi *et al.*, 2002)) ont été obtenues entre le 31 octobre et le 6 décembre 2002. Le champs mesure 21×21.6

minutes carrées. Il inclut le champ GOODS⁶ (voir [Giavalisco et al. \(2004\)](#)) dont les images obtenues avec la caméra ACS (pour *Advanced Camera for Surveys*) du *Hubble Space Telescope* (HST) sont publiques. Ce champ est aussi observé avec les satellites Chandra, XMM et Spitzer. La figure 2.18 représente la zone observée du CDFS.

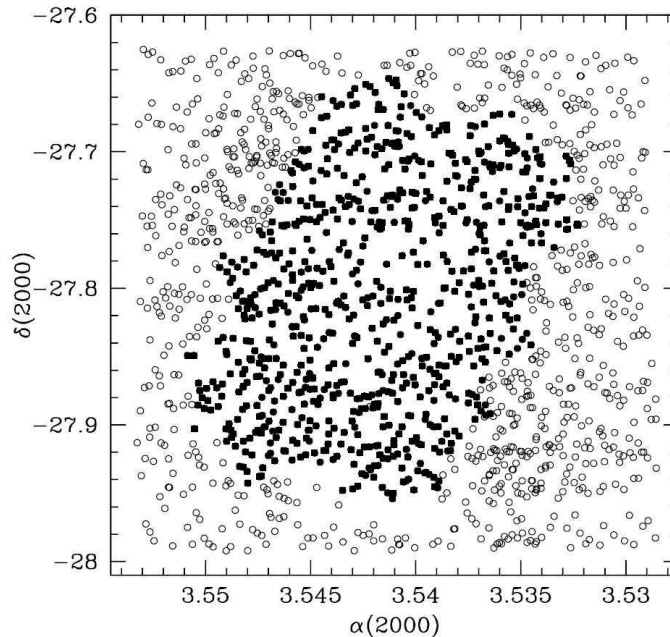


FIG. 2.18 – Objets observés avec VIMOS dans le *Chandra Deep Field South*. Les points noirs sont les objets VVDS observés dans le champ HST-ACS GOODS. Figure extraite de [Le Fèvre et al. \(2004b\)](#)

Cinq pointés VIMOS ont été définis. La sélection des objets s'est faite sur la base du catalogue photométrique du sondage EIS ([Arnouts et al., 2001](#), *ESO Imaging Survey*). Un échantillon complet jusqu'à magnitude $I_{AB} = 24$ de 1599 *redshifts* a ainsi été rendu public au printemps 2004 ([Le Fèvre et al., 2004b](#)). Les données sont disponibles pour toute la communauté internationale via la base de donnée CENCOS⁷ (CENTre de COSmologie).

L'échantillon regroupe 1451 galaxies, 139 étoiles et 10 quasars (table 2.6). Il correspond à 30% du nombre total de galaxies visibles dans le champ jusqu'à à $I_{AB} = 24$. Près de la moitié de cet échantillon, soit 748 *redshifts*, correspond à des objets du champ GOODS. Le taux de complétude, rapport entre le nombre de *redshifts* et le nombre d'objets observés (après avoir déduits du comptage les objets présentant un problème instrumental), est de 88% et monte à 95% en incluant les notes (partie 2.3.2.3) 1 et 11. Le *redshift* moyen est de 0.73.

La précision des *redshifts* a été évaluée de plusieurs manières. Tout d'abord, 160 objets ont été observés deux fois de façon indépendante. La dispersion en vitesse obtenue est de $\sigma_z = 0.0012$ soit 360 km.s^{-1} . Enfin, la comparaison avec les observations spectroscopiques des sondages K20 ([Cimatti et al., 2002](#)) et GOODS indique des dispersions du même ordre (respectivement $\sigma_z = 0.0017$ et $\sigma_z = 0.0013$).

⁶voir <http://www.eso.org/science/goods/>

⁷voir <http://cencosw.oamp.fr/>

	galaxies		étoiles		quasars		non identifié
note	1	2-9	1	2-4	11	12-19	0
nombre	137	1258	3	128	1	9	102
note	21	22-29	21	22-24	211	212-219	20
nombre	10	46	0	8	0	0	20

TAB. 2.6 – Répartition du nombre d'objets en fonction de leur type et de leur note dans le VVDS-CDFS

Enfin, une particularité de ce champ est de présenter de nombreux pics dans sa distribution en *redshifts* (figure 2.19). Les deux principaux sont situés à des *redshifts* moyens de $z=0.667$ et 0.735 et contiennent respectivement 149 et 116 galaxies. Ces groupes sont distribués sur tout le champ, représentant en général d'avantage des murs que des amas bien localisés (Adami *et al.*, 2005).

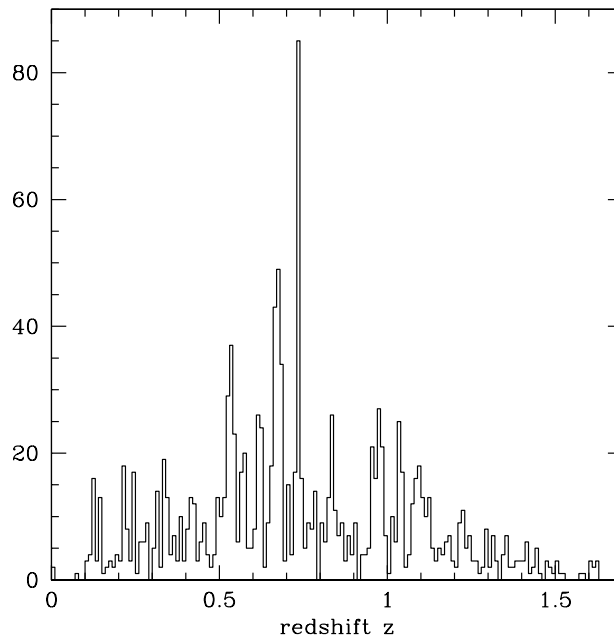


FIG. 2.19 – Distribution en *redshift* du VVDS-CDFS ($17.5 \leq I_{AB} \leq 24$, notes 1 à 9) où l'on distingue les deux murs à $z=0.667$ et 0.735

2.4.2 Les champs *Wide*

Les objets observés dans les champs *Wide* sont plus brillants (magnitude $I_{AB} \leq 22.5$). Les temps de poses sont donc réduits : 10 poses de 5 minutes par pointé soit 50 minutes d'intégration.

Quatre-vingt dix-neuf pointés *Wide* ont été réalisés (septembre 2005), respectivement 25, 21 et 53 sur les champs VVDS-10h, VVDS-14h et VVDS-22h. 6038 objets ont été

observés dans le champ VVDS-10h, 8467 dans le VVDS-14h et 21754 dans le VVDS-22h. Les surfaces couvertes sont respectivement de ~ 1 , ~ 1.43 et ~ 3 degrés carrés. Les trois champs sont représentés figures 2.20 et 2.21. Uniquement un seul passage a été fait. On distingue ainsi clairement la géométrie de l'instrument avec ces 4 quadrants séparés les uns des autres de 2 minutes d'arc (figure 2.3). L'élargissement des zones observées ainsi qu'un deuxième passage permettant de doubler l'échantillonnage spatial sont en cours.

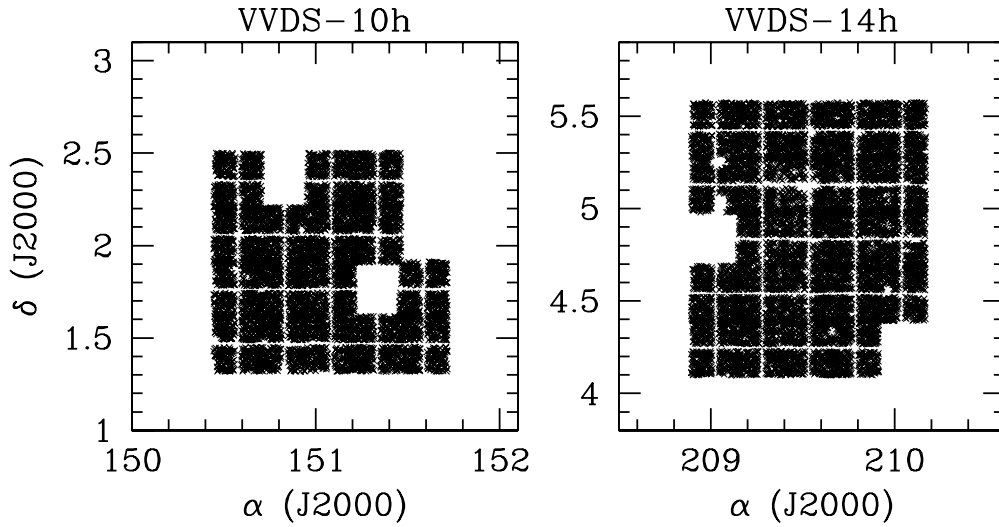


FIG. 2.20 – Objets observés avec VIMOS dans les champs *Wide* VVDS-10h (6038 objets) et VVDS-14h (8467 objets).

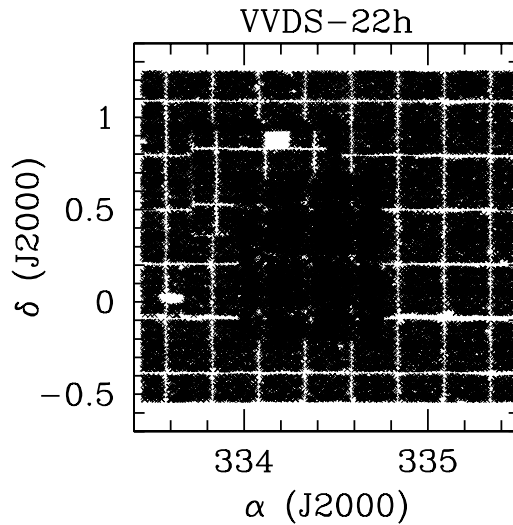


FIG. 2.21 – Objets observés avec VIMOS dans les champs *Wide* VVDS-22h (21754 objets)

Les distributions en *redshifts* des champs VVDS-10h et VVDS-22h sont reportées figure 2.22 (seulement 1908 et 5348 *redshifts* (notes 1 à 9) étaient disponible sur la base

de donnée VVDS début septembre 2005 dans les champs VVDS-10h et VVDS-22h respectivement). Si l'on rassemble ces deux échantillons, le *redshift* moyen des notes de confiance 1 à 9 est de 0.43. Il descend à 0.37 si l'on ne garde que les objets de notes comprises entre 2 et 9. Une description plus détaillée des premières données *Wide* est en préparation ([Garilli et l'équipe VVDS, 2005](#)).

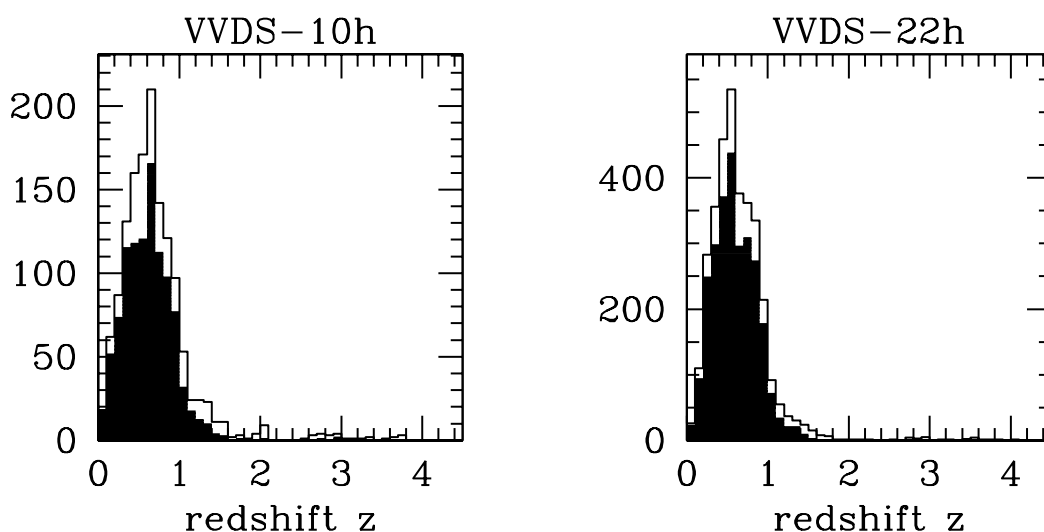


FIG. 2.22 – Distributions en *redshifts* des champs VVDS-10h et VVDS-22h. En noir galaxies de notes de confiance comprise entre 2 et 9, en blanc, notes comprises entre 1 et 9. Figure réalisée à partir des données spectroscopiques disponibles dans ces champs sur la base de données VVDS début septembre 2005

Chapitre 3

Agrégation des galaxies : outil de mesure et erreurs associées

Résumé

3.1	La fonction de corrélation spatiale	81
3.2	Contraintes observationnelles propres au VVDS	85
3.3	$\xi(r_p, \pi)$: l'outil de mesure	87
3.4	Mesure des paramètres r_0 et γ	93

Ce chapitre présente la fonction de corrélation spatiale $\xi(r_p, \pi)$ et le code informatique utilisé dans le cadre du VVDS pour la calculer. Je décris ensuite les contraintes observationnelles propres au VVDS qui ont dû être prises en compte. J'explique enfin la méthode et les outils que j'ai développés afin de mesurer les paramètres r_0 et γ de la fonction de corrélation spatiale $\xi(r)$ ainsi que leurs erreurs associées.

3.1 La fonction de corrélation spatiale

3.1.1 $\xi(r_p, \pi)$ à 2 dimensions

La distribution des galaxies dans l'espace réel est différente de celle observée dans l'espace des *redshifts*. Dans ce dernier, le flot de Hubble crée une distorsion visible aussi bien dans les amas de galaxies que dans des structures plus importantes encore. Pour des galaxies présentant une faible séparation angulaire, la distribution de l'ensemble est étirée selon l'axe du *redshift* par les mouvements aléatoires propres de chaque galaxie. Cet effet est dénommé dans la littérature *Finger-of-God*. Ainsi, la fonction de corrélation $\xi(s)$ dans l'espace des *redshifts* ne présente aucun signal à petite échelle, à cause de cette élongation des structures selon la ligne de visée. Dans le même temps, à grandes échelles, les mouvements d'ensemble des structures les unes vers les autres, engendrés par la gravité, compressent le signal de $\xi(s)$ (Kaiser, 1987; Fisher et al., 1994).

Ces effets peuvent être vus dans l'espace des *redshifts* en développant la fonction de corrélation spatiale selon la ligne de visée (π) et sa composante perpendiculaire (r_p). On définit ainsi la fonction de corrélation $\xi(r_p, \pi)$ dont un exemple est représenté en figure 3.1.

On voit clairement apparaître le *Finger of God* pour de faibles valeurs de r_p . Le signal est très étiré jusqu'à de grandes valeurs de π , traduisant la forte dispersion locale des vitesses propres des galaxies. À grande échelle, le signal est créé par la dynamique globale des structures les unes par rapport aux autres. La fonction de corrélation $\xi(r_p, \pi)$ est le résultat de la convolution de la fonction de corrélation spatiale $\xi(r)$ par la distribution des vitesses propres relatives des galaxies (Peebles, 1980, section 76).

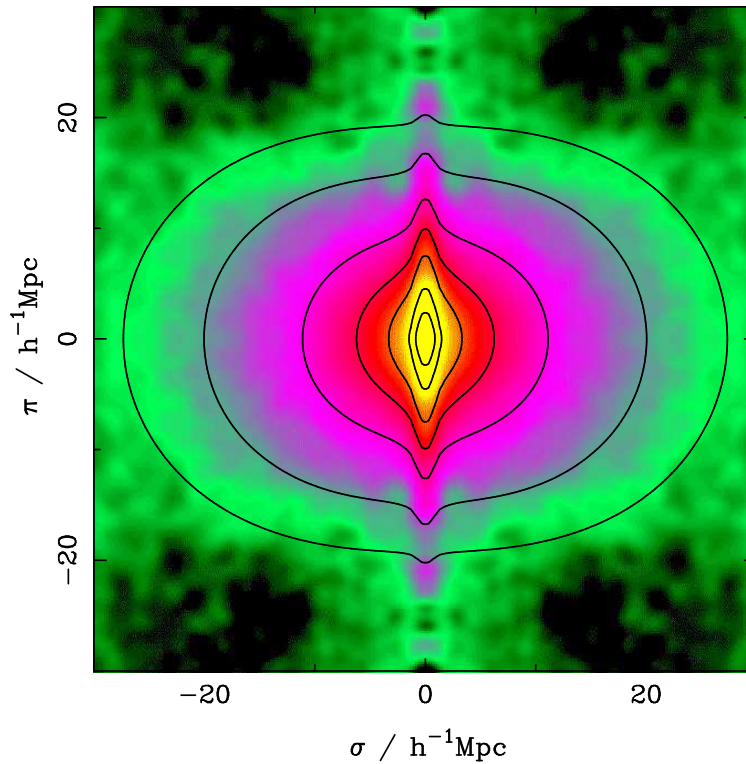


FIG. 3.1 – Exemple de fonction $\xi(r_p, \pi)$ (Peacock *et al.*, 2001, 2dFGRS) montrant la distorsion de l'espace des *redshifts*. Les structures étirées créent le *Finger of God* à petite échelle alors que les mouvements d'ensemble engendrés par la seule gravitation aplatissent le signal à grande échelle (Kaiser, 1987).

Pour mesurer la fonction de corrélation à deux points $\xi(r_p, \pi)$, il faut donc décomposer chaque paire de galaxies transversalement et longitudinalement. Considérons deux galaxies (notées 1 et 2 dans la suite) vues sous un angle θ et identifiées par les vecteurs \vec{v}_1 et \vec{v}_2 représentés sur la figure 3.2. On définit ensuite (Fisher *et al.*, 1994) le vecteur de séparation dans l'espace des *redshifts* \vec{s} (équation 3.1) ainsi que la ligne de visée \vec{l} (équation 3.2)

$$\vec{s} = \vec{v}_1 - \vec{v}_2 \quad (3.1)$$

$$\vec{l} = \frac{1}{2}(\vec{v}_1 + \vec{v}_2) \quad (3.2)$$

Les séparations parallèles (π) et transverses (r_p) sont alors définies par les équations 3.3 et 3.4

$$\pi = \frac{\vec{s} \cdot \vec{l}}{|\vec{l}|} \quad (3.3)$$

$$r_p^2 = \vec{s} \cdot \vec{s} - \pi^2 \quad (3.4)$$

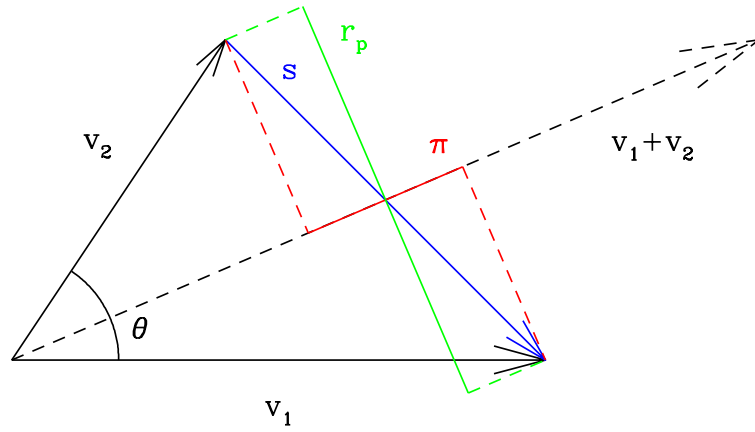


FIG. 3.2 – Deux galaxies, identifiées par l'extrémité des vecteurs \vec{v}_1 et \vec{v}_2 sont vues sous un angle θ . On définit la séparation radiale π et transverse r_p comme respectivement la projection du vecteur \vec{s} sur l'axe de la ligne de visée $\vec{v}_1 + \vec{v}_2$ et sa perpendiculaire.

Si l'on développe un peu le calcul, l'expression de r_p peut s'exprimer uniquement en fonction de l'angle θ sous lequel sont vues les deux galaxies, ainsi que leur position dans l'espace des *redshifts*. On aboutit à l'équation 3.8.

$$r_p^2 = (\vec{v}_1 - \vec{v}_2) \cdot (\vec{v}_1 - \vec{v}_2) - \left((\vec{v}_1 - \vec{v}_2) \cdot \frac{(\vec{v}_1 + \vec{v}_2)}{|\vec{v}_1 + \vec{v}_2|} \right)^2 \quad (3.5)$$

Après développement, réduction au même dénominateur et simplification, on obtient :

$$r_p^2 = \frac{4v_1^2 v_2^2 \sin(\theta)^2}{|\vec{v}_1 + \vec{v}_2|^2} \quad (3.6)$$

$$r_p = \frac{2v_1 v_2}{|\vec{v}_1 + \vec{v}_2|} \cdot \frac{2 \tan\left(\frac{\theta}{2}\right)}{1 + \tan^2\left(\frac{\theta}{2}\right)} \quad (3.7)$$

Si $\theta \ll 1$ alors, d'une part les vecteurs \vec{v}_1 et \vec{v}_2 peuvent être considérés comme étant alignés et d'autre part $\tan\left(\frac{\theta}{2}\right)^2 \ll 1$. Ainsi,

$$r_p \sim \frac{4v_1 v_2}{v_1 + v_2} \tan\left(\frac{\theta}{2}\right) \quad (3.8)$$

Le calcul de la séparation angulaire θ entre les deux galaxies utilise la troisième formule de Gauss (équation 3.9) à partir des coordonnées angulaires (α_1, δ_1) et (α_2, δ_2) de chaque galaxie.

$$\cos(\theta) = \sin(\delta_1) \sin(\delta_2) + \cos(\delta_1) \cos(\delta_2) (\cos(\alpha_1) \cos(\alpha_2) + \sin(\alpha_1) \sin(\alpha_2)) \quad (3.9)$$

Enfin, dans l'espace des *redshifts*, la distance de séparation s s'exprime elle aussi en fonction de l'angle de séparation et des distances qui séparent l'observateur des deux galaxies (équation 3.10).

$$s = \sqrt{v_1^2 + v_2^2 - 2v_1v_2 \cos(\theta)} \quad (3.10)$$

3.1.2 La projection $w_p(r_p)$

Pour s'affranchir de la distorsion de $\xi(r_p, \pi)$ dans l'espace des *redshifts*, on définit sa projection, $w_p(r_p)$, selon la ligne de visée afin de retrouver la fonction de corrélation spatiale $\xi(r)$ dans l'espace réel.

$$w_p(r_p) = 2 \int_0^\infty \xi(r_p, \pi) d\pi \quad (3.11)$$

$$= 2 \int_0^\infty \xi \left[(r_p^2 + y^2)^{1/2} \right] dy \quad (3.12)$$

La fonction que l'on somme dans l'équation 3.12 est la fonction de corrélation dans l'espace réel. Si l'on modélise $\xi(r)$ sous la forme d'une loi de puissance, alors cette intégrale peut être calculée analytiquement et conduit à

$$w_p(r_p) = r_p \left(\frac{r_0}{r_p} \right)^\gamma \times \frac{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{\gamma-1}{2}\right)}{\Gamma\left(\frac{\gamma}{2}\right)} \quad (3.13)$$

où $\Gamma(s)$ est l'usuel fonction Gamma définie sur le demi-plan $\{s \in \mathbb{C} \text{ tel que } \Re(s) > 0\}$ par $\Gamma(s) = \int_0^{+\infty} t^{s-1} e^{-t} dt$.

Cette fonction $w_p(r_p)$ est elle aussi une loi de puissance qui ne dépend que de la longueur de corrélation r_0 et de la pente γ de la fonction de corrélation spatiale $\xi(r)$. L'étude de $w_p(r_p)$ permet donc de mesurer les propriétés d'agrégation d'un échantillon de galaxies en remontant aux valeurs de r_0 et γ .

La limite d'intégration supérieure de $\xi(r_p, \pi)$ doit être définie correctement car elle ne peut être égale à l'infini (équation 3.11). Elle doit être suffisamment grande de façon à mesurer précisément $w_p(r_p)$ sans toutefois intégrer trop de bruit du signal de $\xi(r_p, \pi)$ pour de grandes valeurs de π .

3.2 Contraintes observationnelles propres au VVDS

3.2.1 Les masques photométriques

La géométrie angulaire des différents champs du VVDS est particulière. Ainsi, avant de sélectionner les cibles spectroscopiques, il est nécessaire de définir pour chaque champ un masque photométrique. Ces masques délimitent les zones réellement utilisées. Ils identifient les régions de mauvaise qualité. Cela permet de retirer du catalogue photométrique les régions dégradées par la présence d'une étoile saturée, ou par une mauvaise réduction. La figure 3.3 représente les masques photométriques des champs VVDS-02h, VVDS-CDFS et VVDS-10h. Les régions vides correspondent aux zones retirées du catalogue. La construction d'un de ces masques est complexe et fastidieuse. Le masque photométrique du VVDS-02h a été réalisé par H.J.McCracken, celui du VVDS-10h par Sébastien Foucaud et celui du CDFS par moi-même à partir des contours délimitant les différentes zones à exclure fournis par Stéphane Arnouts. Une version plus restrictive du masque photométrique du VVDS-02h a été réalisée par H.J.McCracken mais entre temps, la sélection des cibles spectroscopiques et leur observation avec VIMOS avait déjà commencé. Toutes les études de l'agrégation des galaxies utilisent donc le premier masque photométrique (figure 3.3 en haut à gauche).

3.2.2 Le critère de sélection observationnelle

Les objets observés en spectroscopie avec VIMOS sont sélectionnés uniquement sur leur magnitude apparente I_{AB} . Celle-ci est comprise entre 17.5 et 24 pour les champs *Deep* et entre 17.5 et 22.5 pour les champs *Wide*. Les objets sélectionnés peuvent donc aussi bien être des étoiles que des galaxies. C'est dans ce catalogue photométrique de référence que vont être sélectionnées les cibles.

3.2.3 La sélection des cibles

Les catalogues photométriques de références du VVDS-02h ou du VVDS-CDFS sont utilisés par le programme SSPOC. À partir des grilles de pointés utilisées sur ces deux champs (figure 2.15 et 3.4), connaissant le rayon apparent de chaque objet, SSPOC va chercher à maximiser le nombre d'objets observable sur chaque pointé. De plus SSPOC garde l'information que tels et tels objets ont déjà été observés afin d'éviter de sélectionner un même objet dans deux pointés VIMOS différents.

3.2.4 L'incertitude sur la mesure du *redshift*

Les observations répétées d'objets des champs VVDS-02h et VVDS-CDFS ont permis d'établir l'erreur de mesure associée aux *redshifts* ($\sigma \sim 275 \text{ km.s}^{-1}$, section 2.4). Cette incertitude sur chaque *redshift* influence la valeur limite d'intégration de $\xi(r_p, \pi)$ que l'on peut utiliser dans le calcul de la projection $w_p(r_p)$. Suite aux différents tests réalisés sur les simulations *GalICS* (partie 4.3), la valeur optimale obtenue comme limite supérieure d'intégration est de $\pi_{\text{max}} = 20 \text{ h}^{-1} \cdot \text{Mpc}$. Cette valeur sera toujours utilisée par la suite lors du calcul de la fonction $w_p(r_p)$.

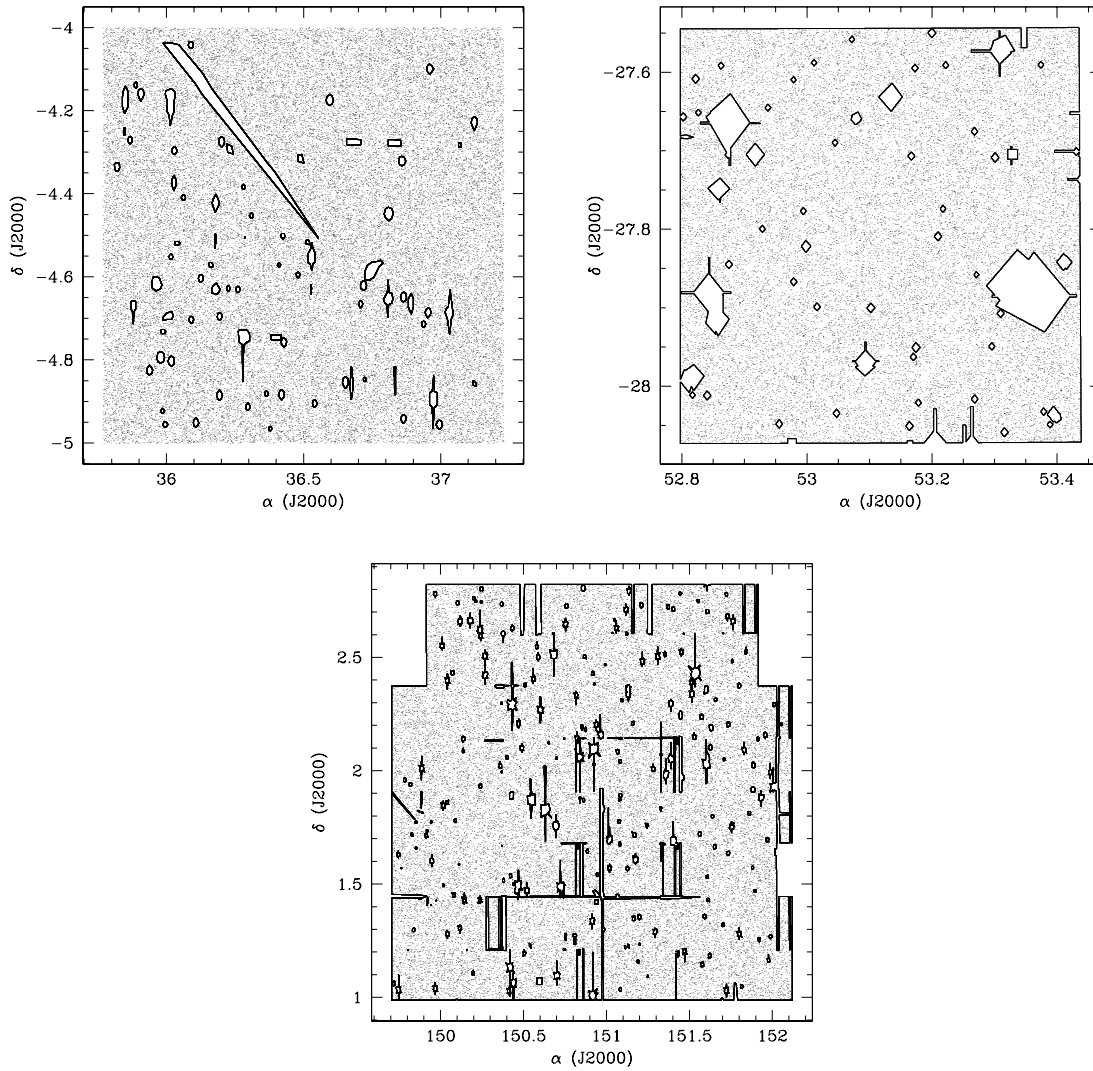


FIG. 3.3 – Masques photométriques du VVDS-02h (en haut à gauche), du VVDS-CDFS (en haut à droite) et du VVDS-F10 utilisés pour la sélection des objets observés avec VIMOS.

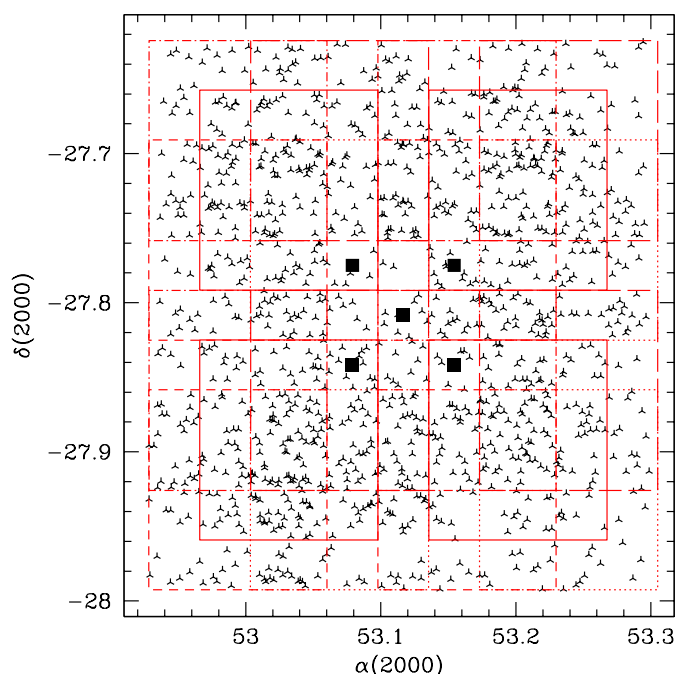


FIG. 3.4 – Grille des 5 pointés VIMOS sur le VVDS-CDFS dont les centres sont représentés par des carrés. Les étoiles correspondent aux objets sélectionnés par SSPOC

3.2.5 L'incomplétude spectroscopique

Ainsi qu'il l'a été détaillé dans la partie 2.4, il est difficile voir impossible de déterminer le *redshift* d'un peu plus de 20% des cibles spectroscopiques. Cette incomplétude doit donc être prise en compte dans nos analyses.

Ce qui nous intéresse est le taux de succès spectroscopique. Celui-ci dépend essentiellement de la magnitude I_{AB} . En effet, il est aisé de comprendre qu'il est plus facile de mesurer le *redshift* d'un objet brillant que faible, car statistiquement, le rapport signal à bruit du spectre de l'objet sera alors meilleurs, et les raies plus facilement identifiables.

Le taux de succès spectroscopique des données du VVDS-02h est reporté page 75 (figure 2.17). Dans la mesure où nous ne considérerons par la suite que les objets dont la mesure du *redshift* est sûre à plus de 75%, seul l'histogramme grisé de la figure 2.17 nous intéresse. Cet histogramme représente la probabilité de mesurer le *redshift* de façon sûre d'une galaxie de magnitude I , sélectionnée par SSPOC. Autrement dit, huit objets sur dix de magnitude $I_{AB} = 23$ observés avec VIMOS ont un *redshift* dont la note de qualité est comprise entre 2 et 9 (section 2.3.2.3).

3.3 $\xi(r_p, \pi)$: l'outil de mesure

Cette partie présente le code utilisé pour calculer la fonction de corrélation à deux dimensions $\xi(r_p, \pi)$. Nous allons décrire son fonctionnement à partir d'un jeu de données : un catalogue photométrique de référence et un catalogue spectroscopique qui en est extrait.

Ce code a été écrit en langage *Fortran 77* par Agnieszka Pollo dans le cadre de son travail post-doctoral à l'*Osservatorio Astronomico di Brera* à Milan-Merate avec Luigi Guzzo. J'ai participé au développement de ce code en réalisant de nombreux tests, et en identifiant et corrigeant certaines erreurs. Il permet de calculer la fonction de corrélation à deux points à deux dimensions $\xi(r_p, \pi)$ ainsi que la fonction de corrélation dans l'espace des *redshifts* $\xi(s)$. L'exécution du programme est configurable grâce à l'utilisation de plusieurs fichiers de paramètres.

J'ai écrit plusieurs programmes en *Fortran 77 et 95* qui permettent de calculer la projection $w_p(r_p)$ et les paramètres r_0 et γ de la fonction de corrélation spatiale. Cette étape (détaillée en partie 3.4) nécessite d'avoir calculé N réalisations différentes de $w_p(r_p)$ (à partir de N simulations différentes par exemple). J'ai donc développé un ensemble de procédures écrites en Shell/Linux afin d'automatiser les différentes étapes et de gérer facilement les données et fichiers de sorties. J'ai aussi fourni un ensemble de routines écrites en *Super Mongo*¹ afin d'exploiter les résultats obtenus. J'ai ainsi rendu disponible à la collaboration VVDS une distribution qui contient le code source de calcul de $\xi(r_p, \pi)$ et de $w_p(r_p)$, les programmes permettant de déterminer r_0 et γ , ainsi que tous les programmes annexes et d'automatisation que j'ai personnellement développé, le tout accompagné d'un manuel utilisateur (annexe A) décrivant l'ensemble des commandes à utiliser.

Je décris ci-après les points principaux du programme de calcul de $\xi(r_p, \pi)$ ainsi que celui de $w_p(r_p)$, des paramètres r_0 , γ et leurs erreurs associées.

3.3.1 Le catalogue de points générés aléatoirement

Pour l'estimation de toute fonction de corrélation, on génère aléatoirement un catalogue de points (partie 1.4.1). Celui-ci doit suivre la géométrie du catalogue de données. La stratégie d'observation du VVDS implique que la densité de galaxies observées spectroscopiquement varie en fonction de la position angulaire. En effet, dans le cas par exemple du champ VVDS-02h, une galaxie de la région centrale a quatre occasions d'être observée puisque quatre quadrants de quatre pointés VIMOS différents se superposent. Par contre, les bords du champ ne sont observés que par deux pointés VIMOS. La densité de galaxies est donc deux fois moins importante sur les bords qu'au centre. De plus, dans certains cas, suite à des problèmes techniques, tel ou tel quadrant de tel pointé VIMOS n'est pas utilisé. Cela diminue donc localement le nombre de galaxies observées en spectroscopie et donc la densité locale.

Le catalogue de points générés aléatoirement doit donc reproduire cette variation de densité. Il est donc généré en suivant la même grille de pointés VIMOS, avec tel ou tel quadrant en moins lorsque c'est le cas, que celle utilisée pour obtenir le catalogue final d'objets spectroscopiquement observés. Ainsi, les coordonnées angulaires sont tirées aléatoirement, pointé VIMOS par pointé VIMOS, quadrant par quadrant. Ne sont conservés que les points pour lesquels les coordonnées ne tombent pas dans une région évidée par la présence d'une étoile saturée ou suite à une mauvaise réduction. On utilise pour cela les masques photométriques définis à la partie 3.2.1.

La partie radiale des points est tirée aléatoirement dans l'intervalle en *redshifts* considéré en suivant la distribution en *redshifts* des données (partie 2.4.1.1, ainsi que l'article

¹voir <http://www.astro.princeton.edu/~rhl/sm/>

de [Le Fèvre et al. \(2005c\)](#)). La distribution radiale alors obtenue correspond à un lissage gaussien de cette distribution en *redshifts*.

Le nombre de points du catalogue aléatoire est défini dès le début. On s'arrange pour que ce nombre soit de plusieurs ordre de grandeur supérieur au nombre de données. Dans la pratique, dans un intervalle en *redshift* Δz de largeur 0.2, on peut compter ~ 1000 galaxies dans le champ VVDS-02h, pour la première époque d'observation, en incluant uniquement les *redshifts* dont la note de qualité est comprise entre 2 et 9. Pour le calcul de $\xi(r_p, \pi)$, jusqu'à 2000 points sont à chaque fois générés aléatoirement pour chaque pointés VIMOS. La grille du champ VVDS-02h compte 20 pointés VIMOS. Le catalogue généré aléatoirement compte donc ~ 40 fois plus d'objets que celui des données.

3.3.2 L'attribution d'un poids à chaque galaxie

De par la stratégie d'observation, on ne connaît pas le *redshift* de toutes les galaxies. En moyenne, on mesure le *redshift* d'une galaxie sur quatre dans le champ VVDS-02h. Il faut donc tenir compte de ce fait lors du calcul de la fonction de corrélation. En effet, quelque soit l'estimateur choisi (équations 1.11 à 1.13), l'échantillonnage spatial joue un rôle important. Il est bien évident que si l'on ne connaît le *redshift* que d'une seule galaxie qui se trouve être dans une région dense de dix galaxies par exemple, le nombre de paires calculées à partir de cette galaxie doit être multiplié par dix. Ainsi, un poids doit être attribué à chaque galaxie lors du comptage des paires.

Différentes méthodes ont été introduites pour corriger des effets d'échantillonnage spatial (*K statistic* d'[Adelberger et al. \(2005\)](#) par exemple). Je présente ci-après les deux approches qui ont été envisagées successivement dans le cadre de l'étude de l'agrégation des galaxies dans le sondage VVDS.

Un poids pour retrouver le nombre d'objets

La première idée est de corriger localement l'échantillonnage spatial, à l'intérieur d'un cercle centré sur un objet du catalogue spectroscopique et dont le rayon est à définir, en estimant le nombre d'objets que l'on aurait dans l'intervalle en *redshift* considéré si toutes les galaxies étaient observées en spectroscopie, avec un taux de succès de mesure de *redshift* de 100%. La figure 3.5 donne un exemple d'une situation réelle. Les triangles représentent les galaxies pour lesquelles le *redshift* a été mesuré et qui sont situées dans l'intervalle d'étude en *redshift* $[z_1 - z_2]$, alors que les galaxies représentées par des carrés en sont exclues. Enfin, les astérisques représentent les objets pour lesquels nous ne disposons d'aucune information sur le *redshift* (car pas observés en spectroscopie, *redshifts* non identifiés ou de faibles niveaux de confiance).

Dans notre exemple, lors du calcul du nombre de paires issues de la galaxie i , il faut tenir compte des 3 galaxies dont les distances sont inconnues. Elles sont peut-être à considérer dans l'intervalle $[z_1 - z_2]$. À ceci est associée une probabilité P pour chacune d'entre elles. Chaque paire issue de la galaxie i doit donc être comptée $(1 + 3 \times P)$. Il ne faut bien évidemment pas tenir compte des autres galaxies présentes localement pour lesquels les *redshifts* z sont connus et tels que $z_1 \leq z \leq z_2$ puisqu'elles seront considérées ultérieurement.

Pour calculer ce poids, on utilise donc l'information angulaire du catalogue photométrique duquel l'échantillon spectroscopique considéré est extrait. Il est important que ces

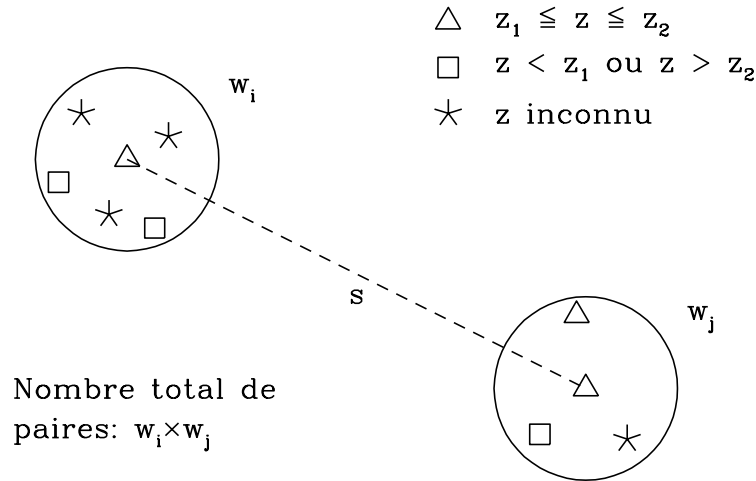


FIG. 3.5 – Un poids w_i est calculé pour chaque galaxie dont on connaît le *redshift* afin de calculer le nombre de paires que l'on aurait si les *redshifts* de toutes les galaxies étaient mesurés

deux catalogues aient les mêmes propriétés de fonction de sélection pour que le poids calculé ait un sens. Ainsi, soit $n_{\text{all}}(i)$ le nombre de galaxies du catalogue photométrique parent inclus dans un cercle de rayon r autour de la galaxie i du catalogue spectroscopique, et $n_{\text{spec}}(i)$ le nombre d'objets dans ce cercle dont le *redshift* est connu, alors le poids w_i associé est

$$w_i = 1 + (n_{\text{all}}(i) - n_{\text{spec}}(i)) \times P \quad (3.14)$$

On peut faire ici l'hypothèse raisonnable que l'échantillon spectroscopique complet fournit la distribution en *redshift* $N(z)$ de tout l'échantillon photométrique de façon non biaisée si moyennée correctement sur une région donnée du ciel. Ainsi, la probabilité $P(z_1 < z_2)$ qu'un objet du catalogue photométrique de référence dont le *redshift* est inconnu se trouve dans l'intervalle $[z_1 - z_2]$ est

$$P(z_1 < z_2) = \frac{\int_{z_1}^{z_2} N(z) dz}{\int_0^\infty N(z) dz} \quad (3.15)$$

Ceci revient en pratique à faire le rapport entre le nombre d'objets du catalogue spectroscopique tels que $z_1 \leq z \leq z_2$ et le nombre total d'objets de ce même catalogue.

Le rayon de la zone sur laquelle est calculée la distribution $N(z)$ est critique. Il doit être assez large pour permettre un échantillonnage correct des structures existantes selon la ligne de visée, tout est restant de taille raisonnable afin de ne pas diluer l'effet d'une structure particulière à un *redshift* donné. En pratique, vue la taille angulaire des champs VVDS-02h et VVDS-CDFS, la distribution en *redshift* utilisée est celle de tout l'échantillon spectroscopique.

Enfin, suite à de nombreux tests, la valeur du rayon du cercle local, dans lequel $n_{\text{all}}(i)$ et $n_{\text{spec}}(i)$ sont calculés, est prise égale à 45 secondes d'arc. Elle est comparable à la taille angulaire d'un spectre VIMOS projeté sur le ciel.

Un poids pour retrouver le nombre de paires

La deuxième méthode pour attribuer un poids à chaque galaxie consiste non pas à retrouver le nombre d'objets manquants mais bien à conserver le nombre de paires locales, où local correspond à l'échelle sous laquelle le problème créé par l'échantillonnage spatial apparaît. En effet, par la méthode précédente, les paires entre deux galaxies proches dont les *redshifts* sont inconnus ne sont pas comptabilisées. Ainsi, localement, chaque paire de galaxies doit être considérée comme

$$w_i \times w_j = \frac{n_{\text{exp}}(i) \times (n_{\text{exp}}(j) - 1)}{n_{\text{obs}}(i) \times (n_{\text{obs}}(j) - 1)} \quad (3.16)$$

où n_{exp} est le nombre estimé d'objets appartenant à l'intervalle en *redshift* $[z_1 - z_2]$ et n_{obs} le nombre d'objets locaux dont le *redshift* appartient à $[z_1 - z_2]$. Ce produit $w_i \times w_j$ correspond donc au nombre de paires estimées normalisé par le nombre de paires présentes dans les données spectroscopiques. La figure 3.6 illustre cette situation. On définit alors un poids à chaque galaxie comme

$$w_i = \sqrt{\frac{n_{\text{exp}}(i) \times (n_{\text{exp}}(i) - 1)}{n_{\text{obs}}(i) \times (n_{\text{obs}}(i) - 1)}} \quad (3.17)$$

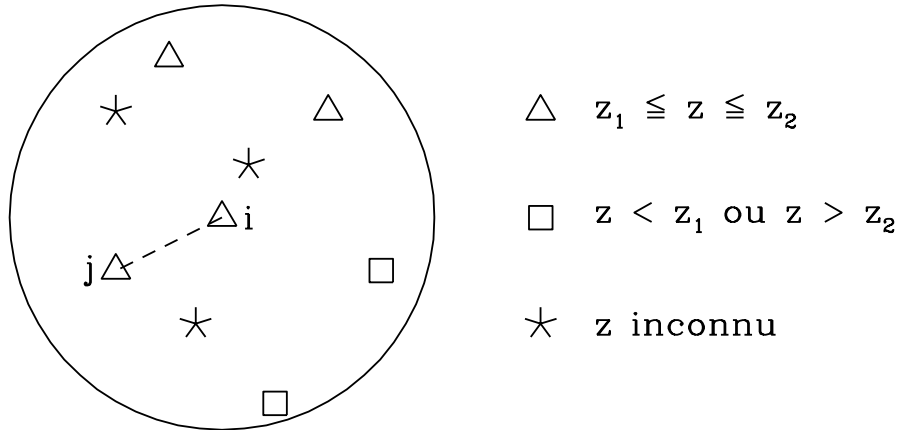


FIG. 3.6 – Le poids w_i est calculé pour chaque galaxie dont on connaît le *redshift* afin de retrouver localement le nombre de paires que l'on aurait si le *redshift* de toutes les galaxies étaient connus.

L'expression de n_{exp} dérive de l'équation 3.14. Avec les mêmes notations jusque là définies, on a

$$n_{\text{exp}}(i) = n_{\text{obs}}(i) + (n_{\text{all}}(i) - n_{\text{spec}}(i)) \times P(z_1 < z_2) \quad (3.18)$$

L'équation 3.17 n'est pas définie si $n_{\text{obs}}(i) = 1$. Dans ce cas, un poids de 1 est alors attribué à la galaxie i .

Pour raffiner, le nombre total d'objet pour lesquels aucune mesure de *redshift* n'est connu doit être multiplié par un facteur f_* qui traduit la probabilité pour un tel objet d'être une galaxie plutôt qu'une étoile. Ce facteur a été évalué à $f_* = 0.92$ à partir des données du VVDS en utilisant les comptages d'étoiles (Le Fèvre *et al.*, 2005c; Pollo *et al.*, 2005).

Cette dernière méthode est celle que nous avons finalement adopté pour attribuer un poids à chaque galaxie du catalogue spectroscopique : un poids pour retrouver le nombre de paires. Le net avantage de cette méthode est qu'elle permet de mieux mesurer le signal de corrélation aux petites échelles ($r_p \lesssim 0.3 h^{-1} \text{Mpc}$)

Le poids attribué aux objets du catalogue généré aléatoirement est toujours égal à l'unité. Il intervient dans le calcul du nombre de paire galaxie-*random* et *random-random*.

3.3.3 Le calcul des distances et l'échantillonnage

La cosmologie dans laquelle les distances sont calculées est déterminée par le choix des valeurs de la densité de matière Ω_m et de celle du vide Ω_Λ . Pour un Univers plat, $\Omega_m + \Omega_\Lambda = 1$.

Dans le cadre d'un modèle d'Univers muni d'une constante cosmologique non nulle, la distance comobile radiale r s'exprime en fonction du *redshift* z , de la vitesse de la lumière c , de la constante de Hubble H_0 et des constantes cosmologiques Ω_m et Ω_Λ (Longair, 1998; Peacock, 1999).

$$r = \frac{c}{H_0} \int_0^z \frac{dz}{\sqrt{(1+z)^2(\Omega_m z + 1) - \Omega_\Lambda z(z+2)}} \quad (3.19)$$

Dans le cas d'un Univers plat, cette expression se simplifie en

$$r = \frac{c}{H_0} \int_0^z \frac{dz}{\sqrt{(1+z)^3 \Omega_m + \Omega_\Lambda}} \quad (3.20)$$

Les densités Ω_m et Ω_Λ sont deux des paramètres d'entrée du code de calcul de $\xi(r_p, \pi)$. La constante de Hubble H_0 est fixée à $100 \text{ km.s}^{-1} \text{Mpc}^{-1}$. Les distances comobiles sont ainsi exprimées en $h^{-1} \text{Mpc}$ (h défini par la relation 1.3).

Une fois mesurée la distance de chaque objet (données et catalogue généré aléatoirement), on calcule les composantes radiales r_p et transverses π de chaque paire galaxie-galaxie, *random-random* et galaxie-*random* comme expliqué dans la partie 3.1.1. On incrémente ensuite au fur et à mesure les nombres de paires radiales $N_\pi(l)$ et transverses $N_{r_p}(k)$ par le produit des poids w_i et w_j , où i et j sont les extrémités (galaxies et/ou *random*) de chaque paire telles que leurs séparation r_p et π vérifient

$$k\Delta r_p \leq r_p \leq (k+1)\Delta r_p \quad (3.21)$$

$$l\Delta \pi \leq \pi \leq (l+1)\Delta \pi \quad (3.22)$$

Les résolutions radiales et transverses sont deux autres paramètres d'entrée du code de calcul. La fonction de corrélation $\xi(r_p, \pi)$ est calculée pour un échantillonnage linéaire des axes r_p et π , pour un échantillonnage logarithmique des deux axes et enfin un échantillonnage logarithmique selon l'axe transverse et linéaire selon l'axe longitudinal.

3.4 Mesure des paramètres r_0 et γ

L'étape précédente a permis de calculer la fonction de corrélation à deux dimensions $\xi(r_p, \pi)$ dans un certain intervalle en *redshift*, à partir d'un jeu de données (catalogue photométrique de référence et catalogue spectroscopique) reproduisant la stratégie d'observation du VVDS. Il peut s'agir de données réelles ou de catalogues fictifs (simulations *GallCS*, partie 4.2). On remonte à la longueur de corrélation r_0 et la pente γ de la fonction de corrélation spatiale $\xi(r)$ en calculant la fonction $w_p(r_p)$ (voir partie 3.1.1) que l'on modélise selon l'équation 3.13. On cherche alors le meilleur couple (r_0, γ) qui permet d'ajuster au mieux la fonction $w_p(r_p)$ mesurée expérimentalement.

Or, il est nécessaire de connaître les erreurs associées à cette dernière, essentiellement liées à la variance cosmique due à la taille des champs observés (voir section 4.1.2). Celles-ci sont évaluées en calculant comme précédemment la fonction de corrélation $\xi(r_p, \pi)$ (puis $w_p(r_p)$) de $N - 1$ autres champs. Dans le cas des simulations *GallCS* (voir partie 4.1.2), nous disposons de 50 champs qui sont par construction quasi indépendants les uns des autres.

J'ai écrit une procédure afin d'automatiser le calcul de la fonction $\xi(r_p, \pi)$ pour chacun des N champs, pour un intervalle de *redshift* donné. Par exemple, le catalogue généré aléatoirement n'est produit qu'une seule fois et réutilisé ensuite. Cette démarche a au préalable dû être validée dans la mesure où la distribution en *redshift* des objets de ce catalogue suit celle des données. Or elle peut changer légèrement d'un champ *GallCS* à un autre. Cette procédure permet donc de simplifier l'utilisation du programme par l'utilisateur lui évitant des manipulations de fichiers pouvant être source d'erreurs. La réutilisation du catalogue généré aléatoirement permet enfin de gagner un temps considérable de calcul dans la mesure où sa production est la partie la plus longue du programme. J'ai ensuite écrit un programme afin de calculer le meilleur couple de paramètres (r_0, γ) de l'une des 50 fonctions $w_p(r_p)$ obtenues, avec leurs erreurs associées.

La méthode utilisée est la suivante. Il s'agit de réaliser un test du χ^2 entre la fonction $w_p(r_p)$ mesurée et le modèle analytique de l'équation 3.13. L'inconvénient est que deux valeurs de w_p à des échelles r_i et r_j différentes sont corrélées. En effet, si l'on s'imagine déplacer dans l'espace une galaxie quelconque, alors tous les comptages de paires galaxie-galaxie et galaxie-*random* sont affectés. Il s'en suit une modification de toutes les valeurs de w_p , ceci à toutes les échelles. Il faut donc réaliser un test du χ^2 généralisé, faisant intervenir la matrice de covariance C de la fonction de corrélation w_p .

On définit la moyenne des valeurs $w_p^k(r_i)$, projection de $\xi(r_p, \pi)$ à la séparation $r_p = r_i$, sur l'ensemble des N champs (indices $k=1$ à N) pour lesquels la fonction de corrélation $\xi(r_p, \pi)$ a été mesurée, par

$$\langle w_p^k(r_i) \rangle_k = \frac{1}{N} \sum_{k=1}^N w_p^k(r_i) \quad (3.23)$$

Le terme (i, j) de la matrice de covariance C est alors par définition

$$C_{ij} = \left\langle \left(w_p^k(r_i) - \langle w_p^k(r_i) \rangle_k \right) \left(w_p^k(r_j) - \langle w_p^k(r_j) \rangle_k \right) \right\rangle_k \quad (3.24)$$

Si les valeurs de w_p n'étaient pas corrélées entre elles, les termes non diagonaux de cette matrice seraient nuls. Dans tous les cas, les termes diagonaux sont égaux à la

variance de w_p à chaque séparation r_p :

$$C_{ii} = \left\langle \left(w_p^k(r_i) - \left\langle w_p^k(r_i) \right\rangle_k \right)^2 \right\rangle_k = \sigma_i^2 \quad (3.25)$$

Après avoir calculé la matrice de covariance, il reste à l'inverser pour calculer la valeur du χ^2 pour un couple (r_0, γ) donné. L'inversion de la matrice C se fait en utilisant la méthode de décomposition de Cholesky (Press *et al.*, 1992, chapitre 2.9) qui exploite les propriétés d'une matrice de covariance à savoir qu'elle est symétrique positive. Soit H son inverse, alors la valeur du χ^2 est donnée par

$$\chi^2 = \sum_{i=1}^{N_p} \sum_{j=1}^{N_p} \left(w_p^{\text{model}}(r_i) - w_p(r_i) \right) H_{ij} \left(w_p^{\text{model}}(r_j) - w_p(r_j) \right) \quad (3.26)$$

avec N_p le nombre de points de mesure d'échantillonnage de $w_p(r_p)$.

Le modèle qui fait intervenir r_0 et γ est celui de l'équation 3.13 et la fonction w_p que l'on cherche justement à modéliser est celle du jeu de données considérées. Le meilleur modèle est celui pour lequel la valeur du χ^2 est minimum. Il est obtenu en parcourant l'espace (r_0, γ) selon une grille. Une autre méthode qui s'appuie sur des considérations géométriques, *the downhill simplex method*, a été testée (algorithme détaillé dans Press *et al.* (1992), chapitre 10.4), et fournit les mêmes résultats mais sans apporter de gains particulier en temps de calcul par exemple.

Nous faisons dans la suite l'hypothèse que les erreurs sur w_p suivent une loi normale. Cette hypothèse est raisonnable dans la mesure où l'échelle sur laquelle nous modélisons w_p (typiquement $\leq 20 h^{-1} \text{Mpc}$) est bien inférieure aux dimensions du sondage. Le théorème de limite central assure alors que la distribution des erreurs sera quasiment gaussienne.

Connaître la valeur du χ^2 permet de tracer les contours d'erreurs à 1, 2 et 3 σ . Ils délimitent les régions de l'espace des paramètres à l'intérieur desquelles une combinaison de ces derniers (pour nous un couple (r_0, γ)) aura un certain degré de confiance, exprimé en pourcentage, d'être exacte.

Le meilleur couple (r_0, γ) est celui qui minimise la valeur du χ^2 , qui vaut alors χ_{\min}^2 . Les régions d'intervalle de confiance sont définies comme celles où la valeur du χ^2 n'excède pas une certaine valeur $\Delta\chi^2$ au-delà de χ_{\min}^2 . La figure 3.7 montre, dans le cas qui nous concerne, un espace de paramètres à deux dimensions, différentes régions pour différentes valeurs de $\Delta\chi^2$. On y a représenté de façon schématique en trait plein les niveaux de confiance joint pour les deux paramètres à 68.3, 90 et 95.4% (ellipses numéro 1, 2 et 3 respectivement).

Dans un espace à deux dimensions, la région de confiance à 68.3% (à 1σ) correspond à celle où $\Delta\chi^2=2.3$. Si l'on s'intéresse à l'incertitude à 1σ associée à l'un des paramètres (r_0 par exemple), il faut alors projeter sur son axe la zone de confiance à deux dimensions obtenue avec la valeur de $\Delta\chi^2$ correspondant à 1σ dans un espace à une dimension, à savoir $\Delta\chi^2=1$. Dans le cas de la figure 3.7, on a représenté ce contour en pointillé. Sa projection sur l'axe de r_0 correspond au segment AB qui représente donc l'intervalle de confiance à 68.3% pour le paramètre r_0 uniquement. La table 3.1 résume les valeurs de $\Delta\chi^2$ et le degré de confiance correspondant pour un et deux degrés de libertés v .

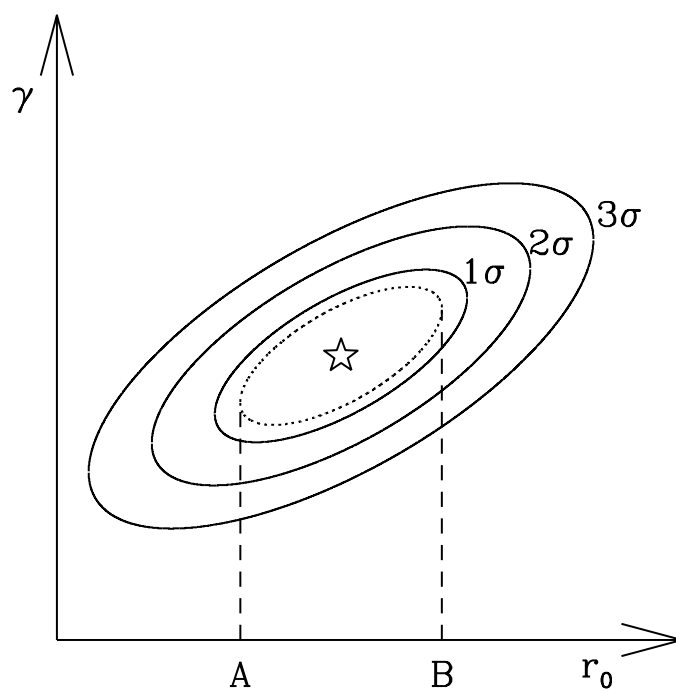


FIG. 3.7 – Niveaux de confiance à 68.3, 90 et 95.4% pour les deux paramètres joints (ellipses en trait plein) et intervalle de confiance à 68.3% pour le seul paramètre r_0 (segment AB).

degré de confiance	ν	
	1	2
68.3%	1.00	2.30
90%	2.71	4.61
95.4%	4.00	6.17

TAB. 3.1 – $\Delta\chi^2$ en fonction du degré de confiance et du nombre de degrés de liberté ν

On a montré dans cette partie la méthode utilisée pour modéliser la projection $w_p(r_p)$ de la fonction de corrélation $\xi(r_p, \pi)$. On obtient ainsi le meilleur couple (r_0, γ) , longueur de corrélation et pente de la fonction de corrélation spatiale $\xi(r)$, avec les niveaux de confiance associés.

Pour obtenir le meilleur couple (r_0, γ) , une méthode de rejection de points de $w_p(r_p)$ à 3σ a été employée. Les points dont l'écart avec le meilleur modèle trouvé est supérieur à 3 fois l'écart-type de w_p à l'échelle considérée sont rejetés un par un. La matrice de covariance perd alors une dimension et est ensuite de nouveau inversée pour recherche de nouveau le modèle s'ajustant le mieux aux données, avec donc un point en moins. Parallèlement, on calcule le meilleur couple (r_0, γ) en ne conservant que les termes diagonaux de la matrice de covariance. On fait ainsi l'hypothèse que les points de mesure de $w_p(r_p)$ ne sont pas corrélés entre eux. On s'attend à ce que le résultat soit peu différent mais avec des erreurs associées plus petites. C'est effectivement ce qui est observé.

Chapitre 4

Calcul des erreurs : utilisation des simulations *GallCS*

Résumé

4.1 Simulations numériques	98
4.2 Création de catalogues fictifs	101
4.3 Comparaison des résultats obtenus en appliquant ou non la stratégie d'observation du VVDS	108
4.4 Estimation des erreurs de mesures	110
4.5 <u>Article I</u> : <i>The VIMOS VLT Deep Survey - Computing the two point correlation statistics and associated uncertainties</i>	112

Ce chapitre présente les simulations d'Univers GallCS. Ces simulations ont été utilisées pour réaliser des catalogues simulés reproduisant la stratégie d'observation du VVDS. Après avoir détaillé la réalisation de ces catalogues, j'explique ensuite comment je les ai utilisés pour tester notre méthode d'analyse des paramètres de la fonction de corrélation spatiale $\xi(r)$ ainsi qu'expliquée dans le chapitre 3 et finalement dériver l'erreur associée à la mesure de la fonction de corrélation.

Le but de ce chapitre est de quantifier les erreurs associées aux mesures des paramètres d'agrégation des galaxies du VVDS. Ces erreurs se décomposent de la façon suivante :

- erreurs liées à la méthode décrite au chapitre 3
- erreurs liées à la taille de l'échantillon observé
- erreurs liées à la variance cosmique

Évaluer les erreurs liées à la méthode revient à vérifier que les paramètres r_0 et γ sont correctement mesurés. Pour cela, j'ai utilisé les simulations d'Univers *GallCS* décrites section 4.1.2. J'ai ainsi pu comparer les valeurs r_0 et γ propres aux simulations à un *redshift* donné à celles mesurées après avoir appliqué toute la stratégie d'observation du VVDS en créant des catalogues de données fictifs (section 4.2). Les résultats sont présentés dans la section 4.3.

Nos catalogues fictifs me permettent ensuite d'estimer en même temps les erreurs liées à la taille de l'échantillon observé et celles à la variance cosmique. En effet, après avoir appliqué la stratégie d'observation et les différentes sources d'incomplétude, le nombre

de galaxies présentes dans ces catalogues, pour un intervalle en *redshift* donné est comparable au VVDS. De plus, quantifier la variance cosmique revient à mesurer la dispersion observée dans les valeurs de r_0 et γ obtenues pour plusieurs champs. Disposer de nombreux champs indépendants par le biais des simulations nous permet donc d'évaluer les erreurs de mesures (section 4.4).

4.1 Simulations numériques

Il devient maintenant possible de réaliser des simulations de volumes d'Univers de plusieurs centaines de méga parsecs comobiles de côté. En 2005, l'équipe de [Springel et al. \(2005\)](#) au Max-Planck Institute für Astrophysik à Garching a construit une simulation de matière noire où $\sim 10^{10}$ particules évoluent dans une boîte dont les dimensions font $500 \text{ h}^{-1} \text{ Mpc}$ de côté, entre *redshift* $z=127$ et $z=0$. Avec 10 fois plus de particules que les précédentes meilleures simulations ([Colberg et al., 2000](#)), la simulation *Millenium* améliore de façon remarquable la résolution spatiale et temporelle, et ceci pour un très grand volume cosmologique. Elle n'a malheureusement été rendue disponible que vers la fin de cette thèse.

Ces nouvelles simulations numériques permettent alors de préparer un relevé fictif du ciel, en imagerie et/ou spectroscopie. Elles permettent d'optimiser les performances d'un sondage profond en testant à l'avance ce que l'on va pouvoir observer. L'intérêt d'utiliser une simulation est aussi de pouvoir tester les programmes de calcul, les outils d'analyse en contrôlant leur bon fonctionnement et en s'assurant qu'ils ne fournissent pas de résultats abérants. Ces simulations permettent aussi de tester la fonction de sélection des sondages ainsi que les biais observationnels attendus.

De plus on dispose directement des propriétés physiques de toutes les galaxies. Cela permet de comparer la mesure d'une quantité physique à partir d'un échantillon de galaxies avec une fonction de sélection donnée à celle obtenue à partir de toutes les galaxies. Ensuite, si l'on dispose de plusieurs simulations d'Univers réalisées indépendamment, il devient possible de quantifier la variance de champ à champ de cette grandeur et d'en déduire des propriétés générales.

4.1.1 Hubble-Volume

Le consortium Virgo¹ a rendu public en 2000 la simulation *Hubble-Volume* ([Frenk et al., 2000](#)). Dans le but d'étudier les modèles d'inflation, deux simulations ΛCDM et τCDM contenant chacune 1 million d'éléments de masse ($2.2 \times 10^{12} M_\odot$) ont été réalisées ([Evrard et al., 2002](#)).

La simulation ΛCDM est une très grosse simulation de matière noire ($\Omega_m=0.3$, $\Omega_\Lambda=0.7$), dont les côtés de la boîte mesurent $3000 \text{ h}^{-1} \text{ Mpc}$. Les 4 catalogues disponibles ont la géométrie soit d'un huitième d'une sphère centrée sur l'observateur jusqu'à *redshift* 1.46, à laquelle on peut ajouter une tranche dont le *redshift* maximum monte à $z=6.8$, soit d'une sphère centrée sur l'observateur jusqu'à *redshift* 0.58, soit d'un cône de 10 degrés par 10 degrés s'étirant jusqu'à $z=4.37$ (figure 4.1). Toutes ces réalisations n'utilisent qu'une seule

¹voir <http://www.virgo.dur.ac.uk/>

boîte d'Univers. Il n'est donc pas nécessaire d'en juxtaposer plusieurs consécutivement pour remplir l'espace. Aucun phénomène de réplication des structures n'apparaît donc.

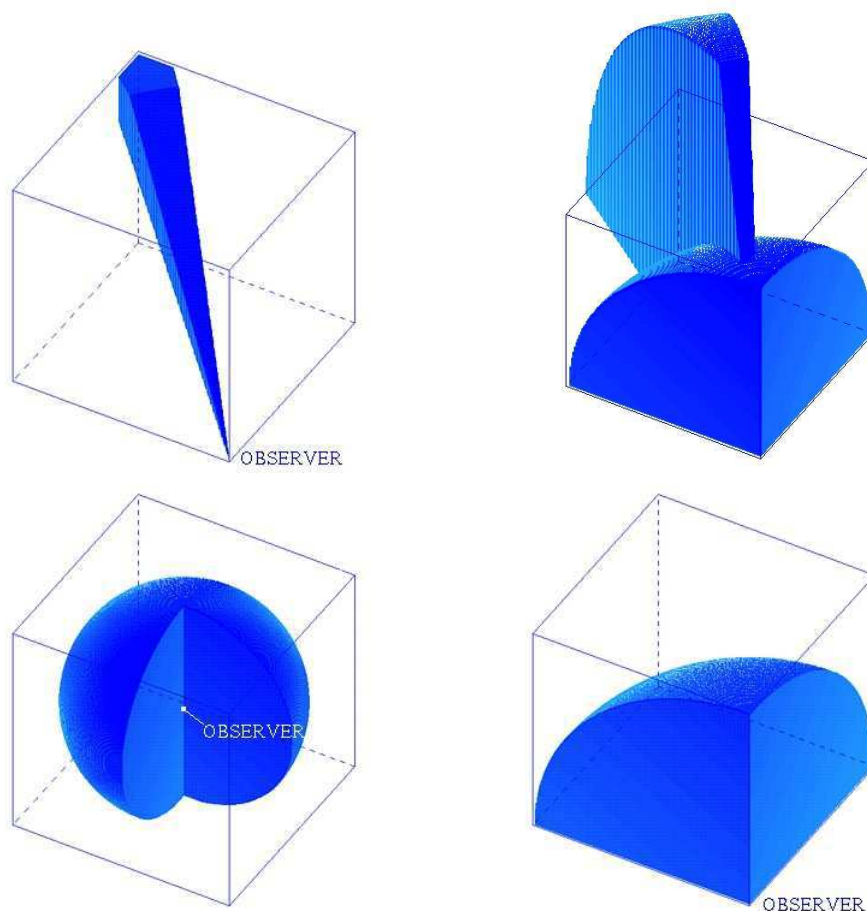


FIG. 4.1 – Géométrie à 3 dimensions des catalogues de matière noire disponibles de la simulation *Hubble-Volume*

Au cours de cette thèse, j'ai commencé par travailler sur le cône du *Hubble-Volume* afin de réaliser des catalogues simulant le champs du VVDS-02h. L'inconvénient principal est que la simulation en question ne contient que des éléments de matière noire (avec position et vitesse). Parallèlement, l'équipe Virgo fournit aussi des informations sur les galaxies en appliquant des recettes semi-analytiques. Il s'agit du projet GIF ([Kauffmann et al., 1999a](#)) qui permet de décrire la formation et l'évolution des galaxies dans une simulation à N corps. Malheureusement pour nous, les catalogues disponibles ne correspondent qu'à des photographies de la simulation à des *redshifts* donnés ($z=0.2, 0.42, 0.62, 0.82, 1.05, 1.46, 2.12, 2.97$). Or, afin de reproduire le sondage VVDS, nous avons besoin d'un catalogue de galaxies simulées dont les *redshifts* augmente de façon continue entre $z=0$ et $z\sim 5$. Donc si l'on en revient au cône de 100 degrés carrés du *Hubble-Volume*, il nous faut appliquer nous même des modèles semi-analytiques afin de décrire les halos de matière noire définis par un nombre minimum de particules de matière de la simulation. Puis il faut

décrire la formation des galaxies dans ces halos, la formation, l'évolution des étoiles et les conséquences sur le contenu en gaz de chaque galaxie, son enrichissement en métaux par l'explosion des étoiles en supernovae, la fusion des galaxies, ...

Ainsi, avant la simulation du *Hubble-Volume*, [Cole et al. \(1998\)](#) ont créé des catalogues adaptés aux sondage du 2dFGRS et SDSS en utilisant 6 modèles différents de biais décrivant les galaxies dans les halos de matière noire puis en appliquant une fonction de luminosité pour décrire les magnitudes absolues des galaxies. Plus récemment, [Coil et al. \(2001\)](#) et [Yan et al. \(2004\)](#) ont utilisé la très grande simulation du *Hubble-Volume* ainsi que la simulation GIF pour préparer le sondage DEEP2. Ils ont alors réalisé des catalogues décrivant la géométrie et fonction de sélection de ce sondage afin de tester leur capacité à analyser la fonction de corrélation à deux points dans ces conditions.

Ainsi, malgré le fait qu'il soit possible de créer 25 champs complètement indépendants de 2 degrés par 2 degrés à partir du cône du *Hubble-Volume*, permettant ainsi de bien quantifier la variance cosmique, le but de cette thèse n'étant pas de développer des modèles semi-analytiques déjà décrits par ailleurs (section 1.6), l'utilisation de la simulation du *Hubble-Volume* a donc été laissée de côté. De même, l'utilisation de la simulation GIF rend difficile la création d'un cône continu en *redshift* puisque ne sont fournis que des instantanés d'Univers à des *redshifts* donnés. Elle a donc elle aussi été écartée. Une solution a été trouvée par l'utilisation de la simulation *GallCS* (section 4.1.2) qui combine, dans le principe, les simulations *Hubble-Volume* et GIF.

4.1.2 *GallCS*

GallCS (pour *Galaxies In Cosmological Simulations*, [Hatton et al. \(2003\)](#)) est un modèle de formation hiérarchique des galaxies qui combine des simulations numériques haute résolution qui décrivent le contenu en matière noire de l'Univers avec des modèles semi-analytiques qui décrivent la matière baryonique.

La simulation à N corps de matière noire utilisée au cours de cette thèse décrit un modèle plat Λ CDM ($h=0.67$, $\Omega_m=0.333$, $\Omega_\Lambda=0.667$). Le volume d'Univers simulé est un cube de côté $L_{\text{box}} = 100 h^{-1} \text{Mpc}$ contenant 256^3 particules de masse $8.272 \times 10^9 M_\odot$. Le spectre de puissance a été initialisé de façon à être en accord avec l'abondance en amas riches observés dans l'Univers local ([Eke et al., 1996](#)). Le champ de densité évolue entre un *redshift* $z=35.59$ et $z=0$, et cent états séparés dans le temps logarithmiquement ont été conservés.

GallCS crée les galaxies à partir de cette simulation en deux étapes. La première consiste à identifier les halos de matière noire constitués d'au moins 20 particules à partir d'un algorithme *friend-to-friend*. L'arbre "généalogique" est construit à partir de l'histoire des particules contenues dans chaque halos. Ensuite, une série de modèles semi-analytiques permettent de faire évoluer la matière baryonique dans ces halos. Ces prescriptions prennent en compte le réchauffement, et refroidissement du gaz dans le halo, la formation des étoiles et les effets sur leur environnement (*feedback*), l'évolution des populations stellaires, leur enrichissement en métaux, la fusion des galaxies et l'instabilité des disques. En gardant la trace de chaque galaxie et ses propriétés à une date donnée, il est ainsi possible d'identifier ses progéniteurs et d'en étudier les propriétés physiques.

L'avantage de *GallCS* est de fournir un large choix de prédictions parmi lesquelles la luminosité (magnitude apparente dans de nombreuses bandes de l'ultraviolet au sub-

millimétrique), des propriétés physiques (dimension des galaxies et de leurs composantes bulbe et disque ainsi que les masses associées par exemple), le *redshift* apparent, le *redshift* cosmologique, ou encore la position des galaxies dans l'espace. De plus [Hatton et al. \(2003\)](#) et [Blaizot et al. \(2004\)](#) ont montré que ces propriétés sont en accord avec de nombreuses observations. Enfin, il est directement possible de créer des catalogues de galaxies avec telles ou telles propriétés à partir de requêtes au format SQL, à travers l'interface du site internet du projet². Ainsi, cinquante champs de 1×1 degré carré sont disponibles et interrogeables sur tel ou tel critère.

On fera souvent référence à ces cinquante champs par la suite (parties 4.2, 4.3). Il est bon avant d'aller plus loin de noter les quelques limitations de ces simulations. Le premier défaut qui nous concerne est la taille (finie) de la boîte de matière noire à l'origine de ces simulations. En effet, pour recréer un cône d'Univers jusqu'à un *redshift* de ~ 5 , il est nécessaire d'aligner plusieurs boîtes le long de la ligne de visée. Ceci entraîne alors des effets de réplication car une même structure peut alors être présente plusieurs fois. De plus, la juxtaposition des boîtes entraîne des problèmes d'effets de bord et de discontinuités. Pour corriger ces effets, [Blaizot et al. \(2005\)](#) réalisent de façon aléatoire des rotations, inversions et translations dans l'espace de la boîte qui a évolué jusqu'au *redshift* z par rapport à celle qui a évolué jusqu'à $z - \delta z$. Le choix de la ligne de visée se fait là encore de façon aléatoire. Ces effets sont résumés par la figure 4.2.

Le deuxième défaut des simulations *GalICS* vient de sa limitation en masse. Toute simulation numérique étant le résultat d'un compromis entre les dimensions de la boîte et le nombre de particules utilisées, il en découle toujours un seuil en masse en-deça duquel les galaxies les moins massives ne sont pas présentes. Les objets les plus faibles sont donc moins bien décrits car moins bien représentés que les objets les plus massifs, les plus lumineux. Dans le cadre de ce travail de thèse, cette limitation n'est pas un problème car les objets du VVDS, limités en magnitude apparente à $I_{AB} = 24$ sont plus brillants que la résolution de *GalICS*.

Enfin, les fluctuations du champ de densité sur des échelles au-delà de $\sim 100 h^{-1} \text{ Mpc}$ ne sont pas présentes dans les simulations, ceci toujours à cause de la taille finie du cube de matière noire utilisé. Ceci implique une mauvaise estimation de la variance cosmique à de telles échelles, ou autrement dit, si le volume sondé est plus grand que la taille du cube. Dans le cas des données du champ VVDS-02h dont les dimensions sont $\sim 0.7 \times 0.7$ degrés carrés, les échelles étudiées à $z \sim 1$ sont de l'ordre de $\sim 30 h^{-1} \text{ Mpc}$, bien inférieures aux dimensions du cube. Les effets de bord du cube se font sentir pour des échelles au-delà de $\sim 10\%$ de la taille du cube ([Blaizot et al., 2005](#)) soit $\sim 10 h^{-1} \text{ Mpc}$.

4.2 Création de catalogues fictifs

La réalisation de catalogues de galaxies reproduisant la stratégie d'observation du VVDS va nous être utile non seulement pour tester notre capacité à mesurer correctement la longueur de corrélation r_0 et la pente γ de la fonction de corrélation spatiale (partie 4.3) mais aussi et surtout à quantifier la variance cosmique de ces paramètres mesurés à partir des données du VVDS (partie 5.1).

Ces catalogues se font en plusieurs étapes :

²voir <http://galics.cosmologie.fr>

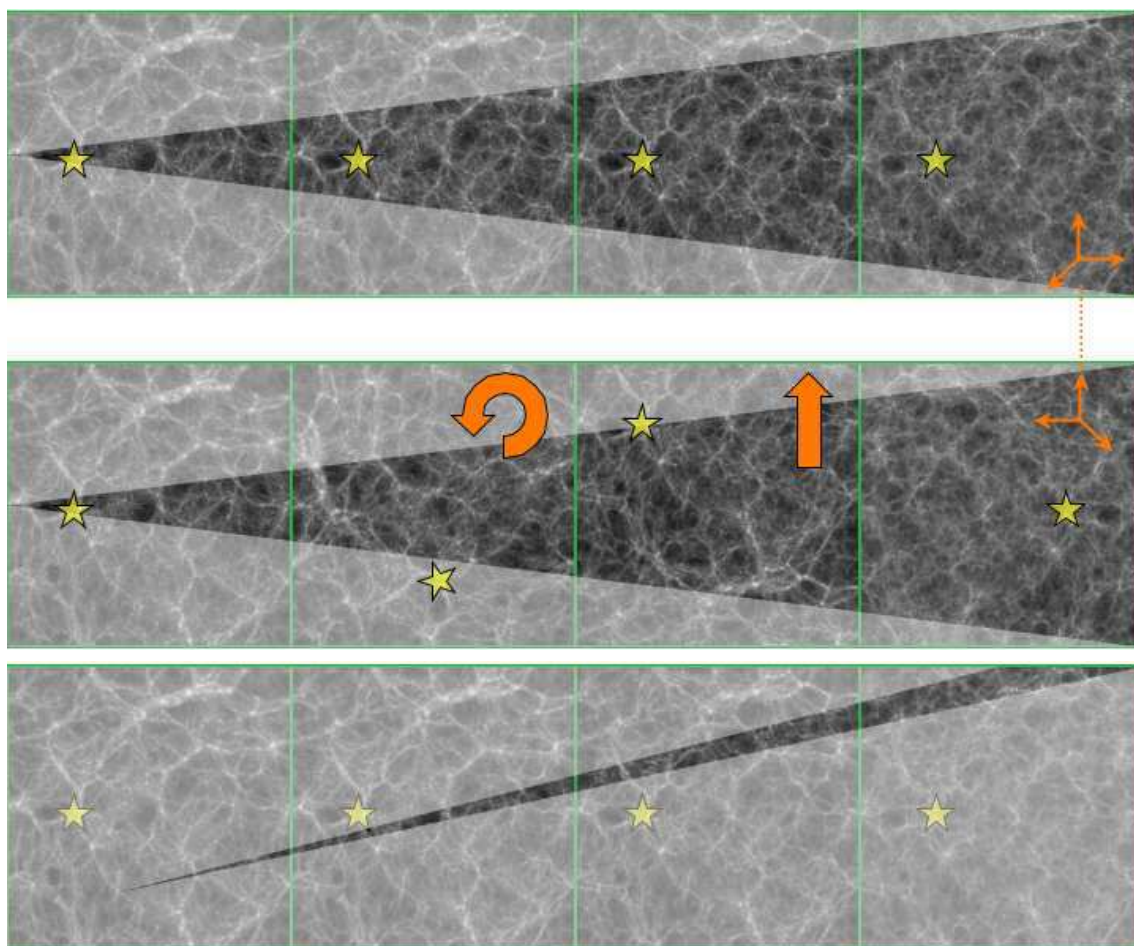


FIG. 4.2 – Processus de réalisation d'un cône *GalICS*. Le panneau du haut montre le résultat d'une juxtaposition directe de plusieurs boîtes sur la même ligne de visée. Dans ce cas, les mêmes structures apparaissent de façon périodique (l'étoile sur la figure). Le panneau du milieu montre l'effet apportées par des transformations géométriques appliquées aléatoirement. De cette manière, deux boîtes consécutives sont décorrélées l'une de l'autre. Le panneau du bas définit enfin une ligne de visée de façon aléatoire. Cette figure est tirée de l'article de [Blaziot et al. \(2005\)](#).

1. ajouter des étoiles
2. définir la géométrie angulaire du champ
3. appliquer la sélection observationnelle
4. tenir compte de l'incertitude sur la mesure des *redshifts*
5. calculer les rayons apparents de chaque objet
6. sélectionner les cibles
7. tenir compte de l'incomplétude spectroscopique

Nous ne détaillons ici que les points 1 et 5, les autres ayant déjà été décrit (partie 3.2).

4.2.1 Les étoiles

Les simulations *GalICS* ne contiennent aucune étoile de la Voie Lactée mais uniquement des galaxies. Or le critère de sélection des objets observés avec VIMOS dans le cadre du VVDS ($I_{AB} = 22.5$ ou 24) fait que des étoiles peuvent être observées. Pour ne pas introduire de biais, il est donc nécessaire d'introduire avant toute chose un nombre réaliste d'étoiles. Il est nécessaire aussi que la distribution en magnitude de celles-ci reproduise celle observée à des latitudes équivalentes. Nous avons donc décidé d'utiliser la simulation développée par Robin *et al.* (2003) de l'Observatoire de Besançon³. Leur modèle produit un catalogue d'étoiles en fonction de différents critères : bande photométrique, coordonnées sur le ciel (la densité d'étoiles change selon que l'on observe dans le plan galactique ou perpendiculairement à celui-ci), . . . Les objets sont tous répartis aux sommets d'une grille de 5×5 qui décrit la zone du ciel demandée. Il est donc nécessaire de redistribuer ces étoiles sur notre champ. Il est très facile de reproduire une telle distribution car celle des étoiles est heureusement aléatoire (la fonction de corrélation des étoiles est nulle à toutes les échelles). Les comptages de ces étoiles, en fonction de la magnitude I_{AB} , reproduit bien celle observée par exemple dans le champ VVDS-02h comme le montre la figure 4.3. L'excès d'étoiles au-delà de $I_{AB} = 20$ est justifiée par la méthode décrite dans McCracken *et al.* (2003). Cette classification tient compte de la compacité des objets. Pour des objets faibles, des galaxies compactes et/ou galaxies à noyaux actifs peuvent être pris pour des étoiles. Cette distribution d'étoiles de Robin *et al.* (2003) est donc bien réaliste. Elle est donc ajoutée aux 50 simulations *GalICS* avant tout autre traitement.

4.2.2 Les rayons apparents de chaque objet

Une fois ces catalogues réalisés, il faut ensuite sélectionner les cibles spectroscopiques. Cette étape se réalise à partir du programme SSPOC (Bottini *et al.*, 2005) utilisé pour la sélection des sources à partir des vraies données VVDS mais elle nécessite la connaissance des rayons apparents des objets. En effet, VMMP-SPOC a besoin de connaître la taille des objets pour pouvoir déterminer la longueur totale de chaque fente. Le programme prend pour cela en compte la nécessité d'avoir suffisamment d'échantillonnage du fond de ciel dans la fente pour pouvoir le soustraire avec une bonne précision.

La base de donnée *GalICS* ne fournit pas cette information directement. Par contre, il est possible d'avoir, pour chaque objet, les longueurs caractéristiques (r_{disc} , r_{bulge} et r_{burst})

³voir <http://bison.obs-besancon.fr/modele/>

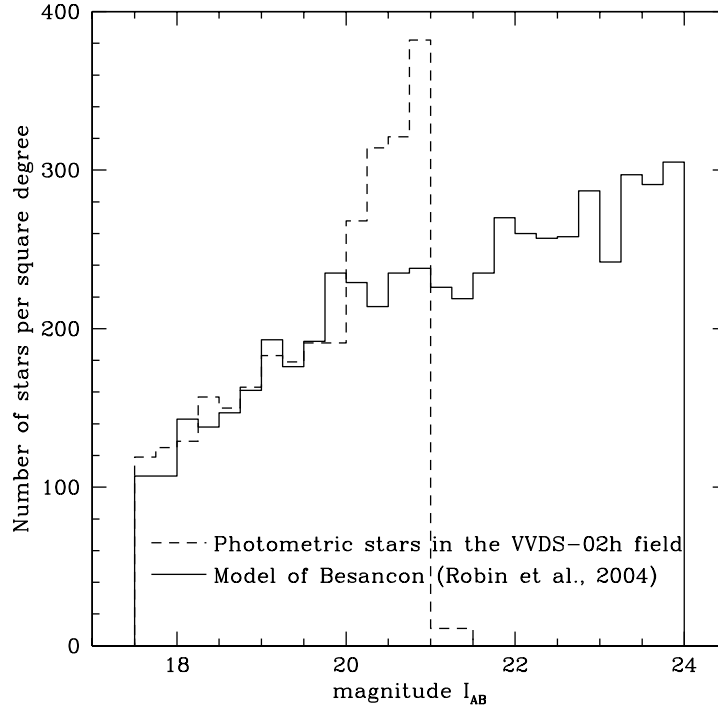


FIG. 4.3 – Distribution des étoiles identifiées comme telles dans le VVDS-02h, comparée à celle ajoutée aux 50 champs *GalICS* à partir du modèle de [Robin et al. \(2003\)](#). L'excès d'étoiles dans le VVDS-02h s'explique par la contamination d'objets compacts puisque cette classification s'appuie sur la compacité des objets plus brillants que $I_{AB} = 21$.

ainsi que les masses totales (M_{disc} , M_{bulge} et M_{burst}) des trois composantes que sont le disque, le bulbe et le *burst*. Le profil de densité du disque suit une loi exponentielle (équation 4.1) alors que les deux autres composantes ont un profil de Hernquist ([Hernquist, 1990](#), équation 4.2).

$$\Sigma(r) = \Sigma_0 \exp\left(-\frac{r}{r_{disc}}\right) \quad (4.1)$$

$$\rho(r) = \frac{M_{total}}{2\pi} \frac{a}{r(a+r)^3} \quad (4.2)$$

Avec l'hypothèse que la luminosité est proportionnelle à la masse, et comme les profils de densité sont définis à l'infini, on considérera le rayon R_{90} apparent qui contient 90% de la masse totale de la galaxie. On calcul ce rayon en tenant compte de l'inclinaison du disque de la galaxie sur la ligne de visée. On utilise alors l'expression du profil de densité du disque ainsi que celle de la brillance de surface des composantes sphériques données par [Hernquist \(1990\)](#). Connaissant la masse totale de la galaxie (fournit dans la base de donnée *GalICS*), on calcule itérativement la valeur de R_{90} .

Ce rayon R_{90} s'exprime en kilo-parsecs. La base de donnée *GalICS* fournit pour chaque

objet sa distance de luminosité D_L ce qui nous permet de convertir R_{90} en un angle θ via la distance angulaire $d\theta$:

$$d\theta = \frac{D_L}{h \times (1+z)^2} \text{ et } \theta = \frac{R_{90}}{d\theta} \quad (4.3)$$

La valeur définitive de l'angle θ est obtenue après convolution par la réponse impulsionnelle de l'imageur à savoir la caméra CFH12k. Enfin, l'angle $\theta/2$, qui s'exprime après conversion en seconde d'arc est discrétisé en un multiple égal à $(n+1/2)$ fois la résolution des images ($0.205''/\text{pixel}$) pour coller à la définition adoptée pour les objets de la base de donnée du VVDS.

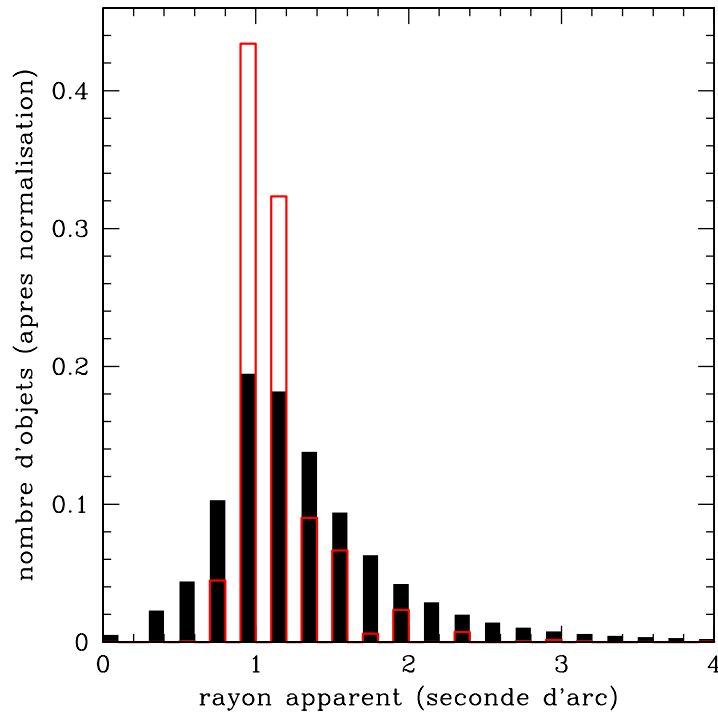


FIG. 4.4 – Distribution moyenne, histogramme plein en noir, des rayons apparents calculés pour les objets des catalogues *GalCS*, comparée à la distribution des rayons apparents des objets du champ VVDS-02h de la base de donnée.

Les rayons apparents des étoiles simulées sont calculés de façon aléatoire, en suivant une distribution de loi gaussienne, dont la moyenne et l'écart type correspondent, par intervalle de magnitude, à ceux mesurés à partir des étoiles de la base de données VVDS jusqu'à $I_{AB} = 20$, puis à partir des étoiles identifiées spectroscopiquement jusqu'à $I_{AB} = 24$.

La figure 4.4 montre la distribution moyenne des rayons apparents calculés pour les objets des catalogues *GalCS*. Elle est directement comparable à celle des objets du champ VVDS-02h de la base de donnée. La distribution obtenue reproduit relativement bien celle des vraies données. Le maximum des deux distributions coïncident. La valeur moyenne du rayon apparent des objets *GalCS* est de 1.33 ± 0.77 secondes d'arc. Celle des objets du VVDS-02h est de 1.11 ± 0.29 .

4.2.3 Comparaison des catalogues - propriétés

On obtient, après toutes les étapes décrites au début de cette partie, plusieurs catalogues pour chacune des 50 simulations *GallCS*. On a ainsi créé des catalogues photométriques sélectionnés en magnitude ($17.5 \leq I_{AB} \leq 24$) qui reproduisent la géométrie des champs du VVDS. À partir de ces catalogues, on applique toutes les étapes conduisant à un échantillon dit spectroscopique. Un dernier catalogue, dit de référence, est aussi créé pour chaque simulation. Il permettra de tester notre capacité à mesurer l'état d'agrégation des galaxies des champs complets à partir uniquement des échantillons spectroscopiques.

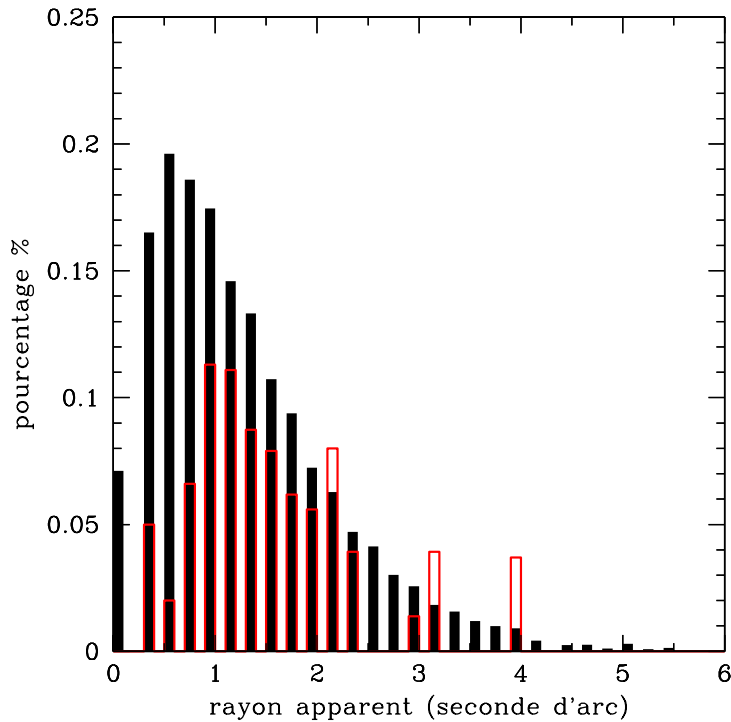


FIG. 4.5 – Pourcentage moyen d'objets sélectionnés par SSPOC en mode optimisé, en fonction de la taille angulaire des objets. En noir, la distribution moyenne obtenue à partir des 50 champs *GallCS*, en rouge, celle des données du VVDS-02h. Dans ce mode, SSPOC a tendance à sélectionner d'avantage les objets de petites tailles, généralement plus faibles.

Il est intéressant ici de regarder quelques propriétés des catalogues *GallCS* obtenus. Les distributions en *redshifts* (figure 4.6) ne sont pas affectées par la stratégie d'observation telle qu'elle a été définie. On pourrait penser que le programme SSPOC, sélectionnant plus facilement des objets faibles, de faibles dimensions angulaires, (comme le montre la figure 4.5) afin de maximiser le nombre de cibles spectroscopiques, puisse avoir une influence sur la distribution en *redshift* des galaxies observées avec VIMOS. Il n'en est rien. La stratégie à plusieurs passages lisse un éventuel effet ayant tendance à sélectionner des objets à très grand *redshift*. La région est alors observée de façon homogène et quasiment aléatoire. Dans le cas du champ VVDS-CDFS, SSPOC est de plus utilisé dans un

mode où il sélectionne aléatoirement les cibles, sans tenir compte de leur taille. Enfin, les comptages de galaxies en fonction de la magnitude I_{AB} sont indiqués figure 4.7.

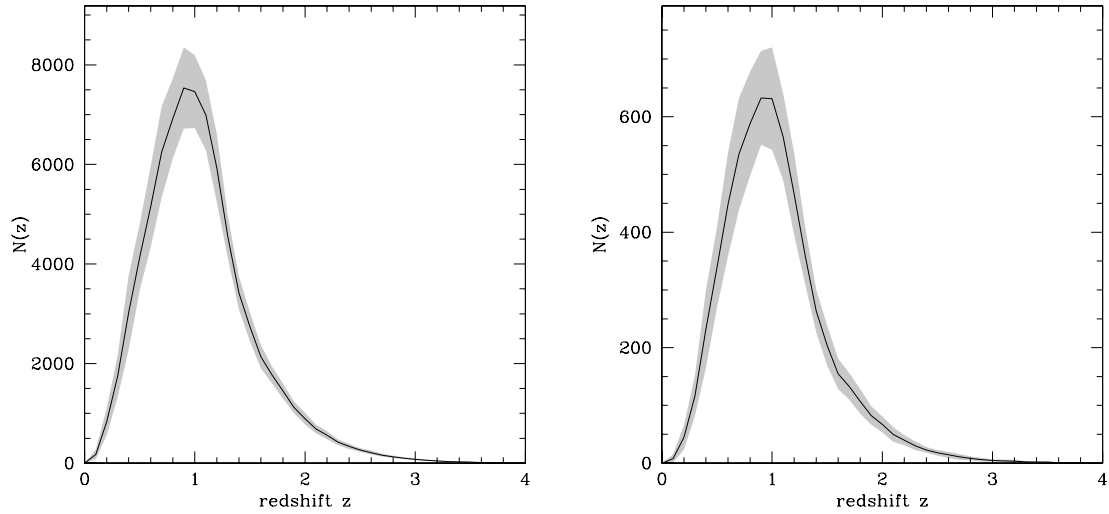


FIG. 4.6 – À gauche, distribution moyenne en *redshift* des 50 champs complets de 1 degré carré de GalICS. À droite, la distribution moyenne en *redshift* des 50 champs complets de 1 degré carré de GalICS après application de la stratégie d'observation du VVDS-02h, jusqu'à l'incomplétude en magnitude. En gris les dispersions à $1-\sigma$ des deux distributions. Si l'on normalise ces distributions par leur nombre d'objets, elles se superposent quasiment.

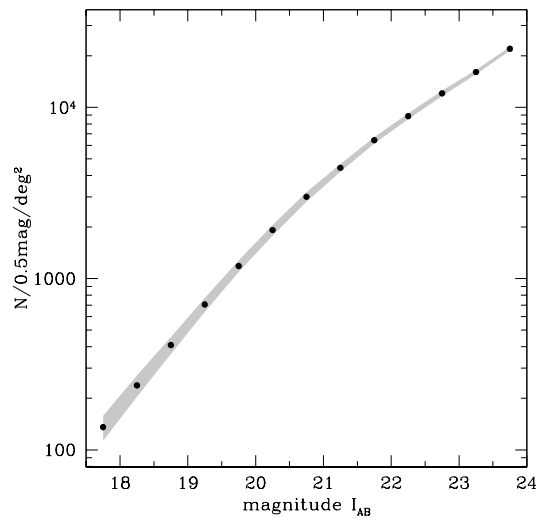


FIG. 4.7 – Comptage moyen en bande I des 50 champs complets de 1 degré carré de GalICS. En gris la dispersion à $1-\sigma$.

4.3 Comparaison des résultats obtenus en appliquant ou non la stratégie d'observation du VVDS

Jusqu'à présent on a présenté la méthode pour mesurer les paramètres de la fonction de corrélation spatiale $\xi(r)$, dans un intervalle de *redshift* donné, pour un jeu de données dont la fonction de sélection est celle de l'un des champs du sondage VVDS (parties 3.3 et 3.4). Mais à ce stade, rien ne dit que nos mesures soient correctes.

L'intérêt d'utiliser des simulations est justement de pouvoir tester nos programmes en mesurant les paramètres réels (r_0, γ) de la fonction de corrélation $\xi(r)$, pour chaque champ *GallCS*, en utilisant toute l'information disponible en position et *redshift* pour toutes les galaxies, et en comparant ces valeurs aux résultats obtenus en appliquant la fonction de sélection du VVDS. Cela permet aussi de tester notre correction de la fonction de sélection en comparant la méthode décrite précédemment avec un code qui calculerait naïvement $\xi(r_p, \pi)$ sans tenir compte de l'échantillonnage spatial qui varie localement.

Une version simplifiée du code de calcul de la fonction de corrélation $\xi(r_p, \pi)$ a été utilisée pour ces tests. Les principes généraux énoncés plus haut restent valide. Simplement, ce code étant appliqué à l'ensemble du catalogue photométrique, il n'est pas nécessaire de construire le catalogue aléatoire en suivant la géométrie des pointés VIMOS. Il est généré uniformément sur tout le champ d'étude. Pour la même raison, le poids attribué à chaque galaxie pour le comptage des paires est mis à 1.

La figure 4.8 montre, pour quatre intervalles en *redshifts* $z=[0.5-0.7]$, $[0.7-0.9]$, $[0.9-1.1]$ et $[1.1-1.3]$, la fonction de corrélation $\xi(r_p, \pi)$ calculée à partir du catalogue obtenu avec la fonction de sélection du champ VVDS-02h pour l'un des 50 champs *GallCS*, en utilisant, à gauche, un algorithme naïf, et, à droite, le code détaillé précédemment (partie 3.3). En pointillé est représentée la fonction $\xi(r_p, \pi)$ mesurée en utilisant les positions angulaires et *redshifts* de toutes les galaxies présentent dans le même champ (en fait la plus petite région rectangulaire qui englobe toute la région définie par les 20 pointés VIMOS). Cette fonction est la référence que l'on cherche à retrouver en utilisant uniquement les données observées spectroscopiquement, avec toutes les incomplétudes déjà décrites. Les lignes de niveau en gras sur les huit quadrants de cette figure tracent $\xi(r_p, \pi)=1$. On constate, à tout *redshift*, un excellent accord en utilisant notre algorithme de correction (catalogue généré aléatoirement, poids attribué à chaque galaxie, ...). On retrouve cet accord avec la figure 4.9 qui montre la projection $w_p(r_p)$ des différentes fonctions $\xi(r_p, \pi)$. Là où un algorithme naïf surestime $w_p(r_p)$ (dans cet exemple tout au moins), notre méthode permet de retrouver la référence.

Ce qui nous intéresse maintenant au premier plan est de quantifier l'erreur sur la longueur de corrélation r_0 et la pente γ . La figure 4.10 montre, pour les quatre intervalles en *redshifts* précédents, les distributions des différences relatives, pour chaque champ *GallCS*, entre la valeur de r_0 (respectivement γ) mesurée connaissant positions et *redshifts* de toutes les galaxies et celle mesurée à partir du sous échantillon obtenu dans les conditions du VVDS.

Le tableau 4.1 rassemble les valeurs moyennes de ces différences relatives pour r_0 et γ . Ces valeurs fluctuent autour de zéro. Aucun effet systématique n'apparaît. La correction apportée par le code pour corriger les effets de la stratégie d'observation du VVDS fonctionne donc correctement puisque les valeurs mesurées sont en accord avec celles attendues.

Ces différents résultats valident donc la méthode. Le code utilisé fonctionne de façon optimale et nous sommes capable de mesurer la longueur de corrélation et la pente de la fonction de corrélation spatiale $\xi(r)$ à moins de 5-10%, valeur significativement plus petite que celle des incertitudes liées à la variance cosmique qui sont de l'ordre de 15-20%.

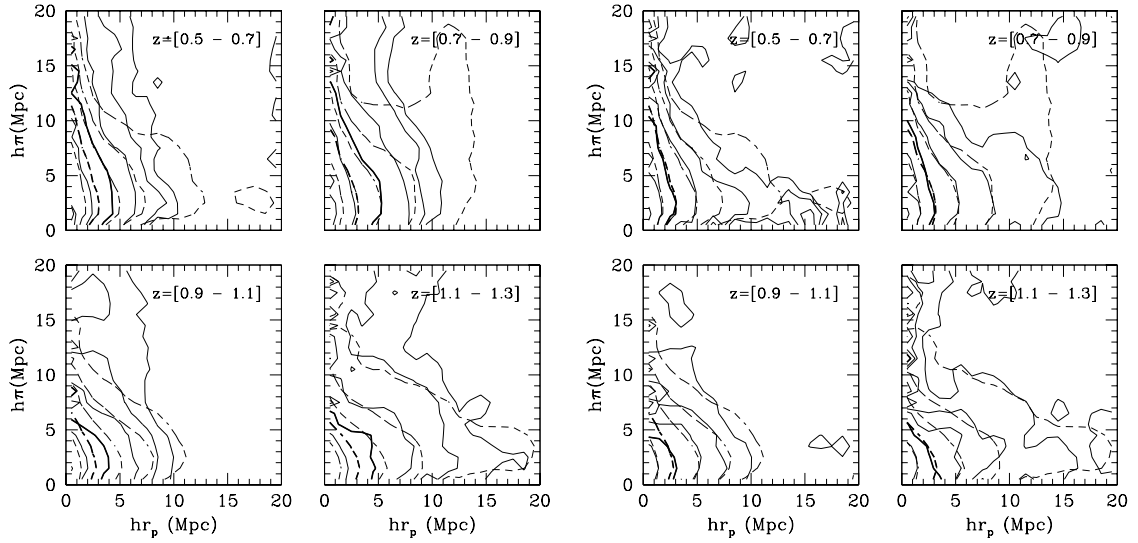


FIG. 4.8 – $\xi(r_p, \pi)$ mesurée dans quatre intervalles en *redshift*. La référence que l'on cherche à retrouver est en pointillé. À gauche, sans correction, à droite avec correction. Figure 9 de Pollo *et al.* (2005)

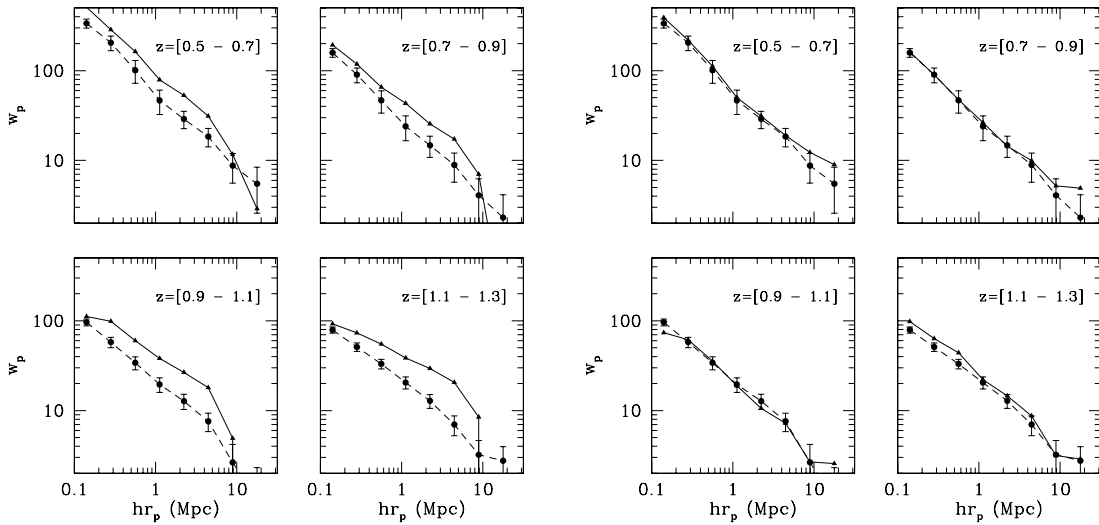


FIG. 4.9 – $w_p(r_p)$ mesurée dans quatre intervalles en *redshift*. La référence que l'on cherche à retrouver est en pointillé. À gauche, sans correction, à droite avec correction. Figure 10 de Pollo *et al.* (2005)

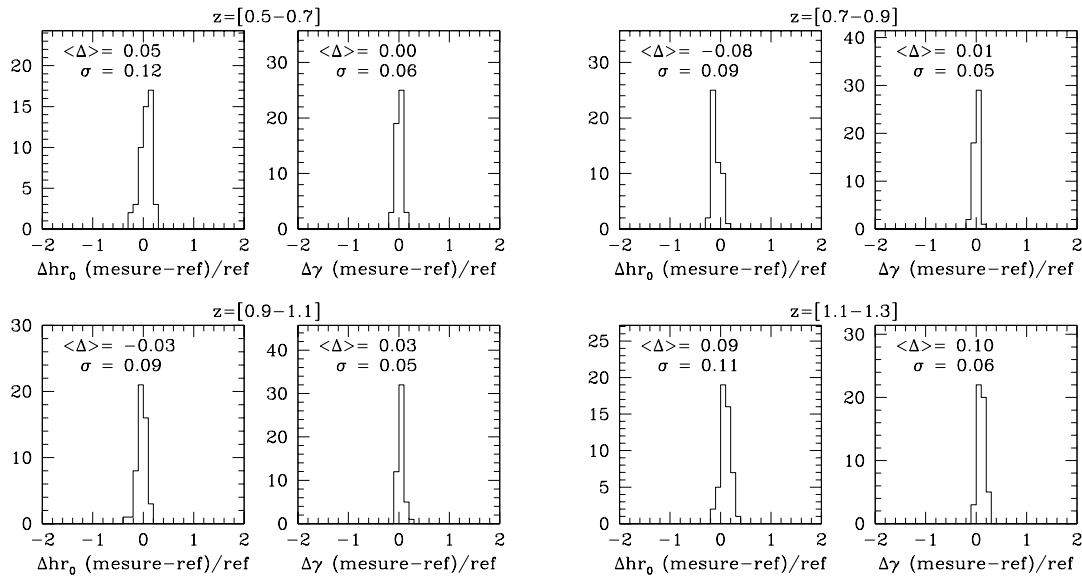


FIG. 4.10 – Distribution des différences relatives, pour chaque champ *GalICS*, entre la valeur de r_0 (respectivement γ) de référence et celle mesurée dans les conditions du VVDS, pour 4 intervalles en *redshifts*.

<i>redshift</i>	erreur relative moyenne	
	r_0	γ
[0.2 – 0.5]	-1%	2%
[0.5 – 0.7]	5%	0%
[0.7 – 0.9]	-8%	1%
[0.9 – 1.1]	-3%	3%
[1.1 – 1.3]	9%	10%
[1.3 – 2.1]	9%	9%

TAB. 4.1 – Écart relatif entre la longueur de corrélation r_0 (respectivement la pente γ) mesurée et la référence.

4.4 Estimation des erreurs de mesures

Le code de calcul et la méthode pour mesurer les paramètres de la fonction de corrélation spatiale étant validés, les simulations *GalICS* vont encore nous servir, cette fois dans le but de dériver les erreurs finales associées à la mesure de $w_p(r_p)$ à partir de l'échantillon de galaxies du VVDS. Par contre il est nécessaire d'ajuster les erreurs calculées à partir des 50 champs *GalICS* à la différence entre l'amplitude de $w_p(r_p)$ mesurée dans le VVDS et celle mesurée dans les simulations. Les propriétés d'agrégation des galaxies de la simulation *GalICS* sont en effet relativement différentes de celles observées dans les données du VVDS.

L'idée est, dans un premier temps, de calculer les paramètres r_0 et γ de la projection $w_p(r_p)$ obtenue à partir des données VVDS dans un intervalle en *redshift* en utilisant une technique de rééchantillonnage *bootstrap* des données du VVDS (Barrow *et al.*, 1984; Mo *et al.*, 1992). Cette méthode consiste en un tirage aléatoire avec remise de N galaxies parmi N . On réalise à chaque fois 100 tirages de la sorte et on calcule la fonction $\xi(r_p, \pi)$ associée. On applique ensuite la méthode de la partie 3.4 à partir de ces 100 réalisations pour mesurer le meilleur couple (r_0, γ) associé aux données du VVDS.

Les erreurs obtenues par cette méthode de rééchantillonnage ont tendance à être sous évaluées comme l'illustre la figure 4.11. On a reporté sur cette figure l'écart relatif entre les erreurs sur $w_p(r_p)$ estimées à partir d'un rééchantillonnage *bootstrap* des données VVDS et celles calculées à partir des simulations *GallCS*, dans le cas du champ VVDS-02h. Le rapport $K(r_p)$ est défini par

$$K(r_p) = \frac{\sigma_{\text{GallCS}} / w_p^{\text{GallCS}}}{\sigma_{\text{bootstrap}} / w_p^{\text{bootstrap}}} \quad (4.4)$$

On ne conserve donc à ce stade que les valeurs de r_0 et γ mesurées par la technique du *bootstrap* sur les données VVDS. Les erreurs réalistes sont calculées à partir des champs *GallCS* sur lesquels la stratégie d'observation a été appliquées. Les 50 champs doivent en effet nous donner une représentation de la variance cosmique. Pour l'intervalle en *redshift* considéré, on calcule donc les fonctions $\xi(r_p, \pi)$ et $w_p(r_p)$. On applique la méthode décrite pour un champ *GallCS* à la moyenne des 50 champs. On obtient ainsi un couple $(r_0, \gamma)_{\text{GallCS}}$ avec des contours d'erreurs associés $\chi^2(r_p, \gamma)$. Il reste ensuite à recentrer de façon relative ces contours d'erreurs autour du couple (r_0, γ) obtenu à partir du rééchantillonnage *bootstrap* des données VVDS :

$$\begin{aligned} r_p &\longleftarrow r_p \times \frac{r_0^{\text{GallCS}}}{r_0^{\text{VVDS}}} \\ \gamma &\longleftarrow \gamma \times \frac{\gamma^{\text{GallCS}}}{\gamma^{\text{VVDS}}} \end{aligned} \quad (4.5)$$

Afin de tenir compte de la différence entre les distributions en *redshifts* des champs *GallCS* et celles des champs du VVDS, la carte des contours d'erreurs $\chi^2(r_p, \gamma)$ obtenue à partir des champs *GallCS* est multipliée par un facteur $N_{\text{VVDS}}/N_{\text{GallCS}}$. Ce facteur est défini à partir de N_{VVDS} , le nombre de galaxies du VVDS dans l'intervalle en *redshifts* considéré et N_{GallCS} le nombre moyen de galaxies sur les 50 champs *GallCS* dans le même intervalle en z .

On obtient ainsi les erreurs associées à la longueur de corrélation r_0 et la pente γ mesurées dans les données du VVDS. Ces erreurs reproduisent la variance de champ à champ observée dans les simulations *GallCS*, normalisée à l'amplitude de corrélation mesurée dans le VVDS, et donc la variance cosmique.

La description de ce travail, de la construction des catalogues fictifs *GallCS* jusqu'à la méthode de détermination des erreurs sur r_0 et γ mesurées dans le VVDS, a été publiée dans la revue *Astronomy & Astrophysics* de septembre 2005. L'article *The VIMOS VLT deep survey : Computing the two point correlation statistics and associated uncertainties* de Pollo *et al.* (2005) est présenté dans la suite de ce chapitre, section 4.5.

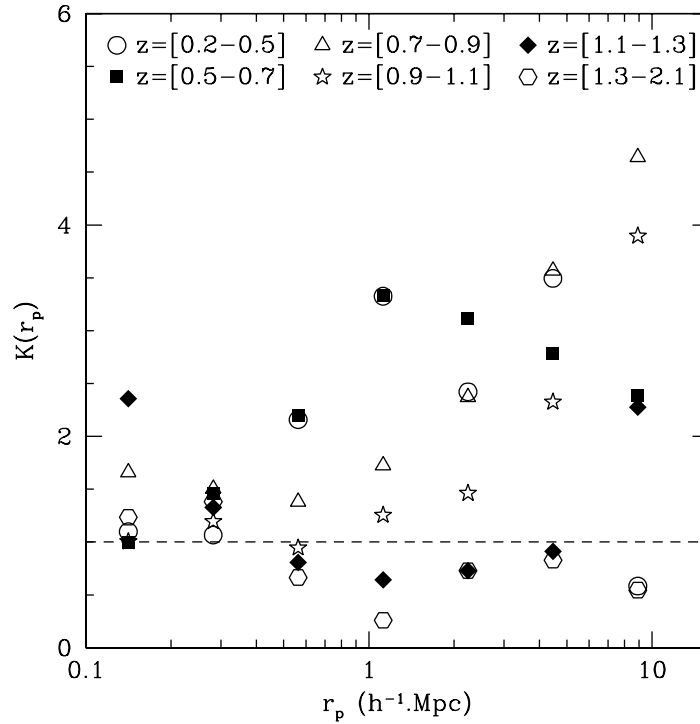


FIG. 4.11 – Différence relative entre les erreurs sur $w_p(r_p)$ estimées à partir d'un rééchantillonnage *bootstrap* des données VVDS et celles calculées à partir des simulations *GalICS*. La valeur de $K(r_p)$ est définie par l'expression 4.4

4.5 Article I : *The VIMOS VLT Deep Survey - Computing the two point correlation statistics and associated uncertainties*

Cet article a été publié dans la revue européenne *Astronomy & Astrophysics* par Pollo, A., Meneux, B., Guzzo, L., Le Fèvre, O. *et al.* (2005, A&A, 439). Il présente l'outil développé et utilisé dans le cadre du VVDS pour calculer la fonction de corrélation spatiale à deux points. La construction des 50 catalogues fictifs *GalICS* y est détaillé ainsi que la méthode permettant de corriger la stratégie d'observation du VVDS, et le calcul de la longueur de corrélation r_0 et la pente γ .

The VIMOS VLT deep survey[★]

Computing the two point correlation statistics and associated uncertainties

A. Pollo¹, B. Meneux², L. Guzzo¹, O. Le Fèvre², J. Blaizot², A. Cappi³, A. Iovino⁴,
 C. Marinoni², H. J. McCracken^{5,6}, D. Bottini⁷, B. Garilli⁷, V. Le Brun², D. Maccagni⁷, J. P. Picat⁸, R. Scaramella⁹,
 M. Scodeggio⁷, L. Tresse², G. Vettolani¹⁰, A. Zanicelli¹⁰, C. Adami², M. Arnaboldi¹³, S. Arnouts², S. Bardelli³,
 M. Bolzonella³, S. Charlot^{5,11}, P. Ciliegi¹⁰, T. Contini⁸, S. Foucaud⁷, P. Franzetti⁷, I. Gavignaud⁸, O. Ilbert²,
 B. Marano¹², G. Mathez⁸, A. Mazure², R. Merighi³, S. Paltani², R. Pellò⁸, L. Pozzetti³, M. Radovich¹³,
 G. Zamorani³, E. Zucca³, M. Bondi¹⁰, A. Bongiorno³, G. Busarello¹³, L. Gregorini¹⁰, F. Lamareille⁸,
 Y. Mellier^{5,6}, P. Merluzzi¹³, V. Ripepi¹³, and D. Rizzo⁸

¹ INAF – Osservatorio Astronomico di Brera, via Bianchi 46, Merate (LC), Italy
 e-mail: apollo@merate.mi.astro.it

² Laboratoire d’Astrophysique de Marseille, UMR 6110 CNRS, Université de Provence, Traverse du Siphon-Les trois Lucs, 13012 Marseille, France

³ INAF – Osservatorio Astronomico di Bologna, via Ranzani 1, 40127 Bologna, Italy

⁴ INAF – Osservatorio Astronomico di Brera, via Brera, 20121 Milan, Italy

⁵ Institut d’Astrophysique de Paris, UMR 7095, 98bis Bd. Arago, 75014 Paris, France

⁶ Observatoire de Paris, LERMA, UMR 8112, 61 Av. de l’Observatoire, 75014 Paris, France

⁷ IASF – INAF, 20133 Milano, Italy

⁸ Laboratoire d’Astrophysique – Observatoire Midi-Pyrénées, 31400 Toulouse, France

⁹ INAF – Osservatorio Astronomico di Roma, Italy

¹⁰ INAF – Istituto di Radio-Astronomia, Bologna, Italy

¹¹ Max-Planck-Institut für Astrophysik, 85741 Garching, Germany

¹² Università di Bologna, Dipartimento di Astronomia, via Ranzani 1, 40127 Bologna, Italy

¹³ INAF – Osservatorio Astronomico di Capodimonte, via Moiariello 16, 80131 Napoli, Italy

Received 6 September 2004 / Accepted 10 March 2005

Abstract. We present a detailed description of the methods used to compute the three-dimensional two-point galaxy correlation function in the VIMOS-VLT deep survey (VVDS). We investigate how instrumental selection effects and observational biases affect the measurements and identify the methods to correct for them. We quantify the accuracy of our corrections using an ensemble of 50 mock galaxy surveys generated with the *GalICS* semi-analytic model of galaxy formation which incorporate the selection biases and tiling strategy of the real data. We demonstrate that we are able to recover the real-space two-point correlation function $\xi(s)$ and the projected correlation function $w_p(r_p)$ to an accuracy better than 10% on scales larger than $1 h^{-1}$ Mpc with the sampling strategy used for the first epoch VVDS data. The large number of simulated surveys allows us to provide a reliable estimate of the cosmic variance on the measurements of the correlation length r_0 at $z \sim 1$, of about 15–20% for the first epoch VVDS observation while any residual systematic effect in the measurements of r_0 is always below 5%. The error estimation and measurement techniques outlined in this paper are being used in several parallel studies which investigate in detail the clustering properties of galaxies in the VVDS.

Key words. cosmology: large scale structure of Universe – methods: statistical – galaxies: evolution – surveys

1. Introduction

The VIMOS VLT Deep Survey (VVDS, Le Fèvre et al. 2005a) is dedicated to study the evolution of galaxies and large scale

structure to $z \sim 2$ with a significant fraction of galaxies reaching $z \sim 4$. The VVDS spectroscopic survey is performed with the VIMOS spectrograph at the European Southern Observatory Very Large Telescope and complemented with multi-color BVRI imaging data obtained at the CFHT telescope (McCracken et al. 2003; Le Fèvre et al. 2004). The complete survey will consist of four fields of 2° by 2° each, with multi-band photometry coverage in the BVRI (and partly UJK) bands.

[★] Based on data obtained with the European Southern Observatory Very Large Telescope, Paranal, Chile, program 070.A-9007(A), and on data obtained at the Canada-France-Hawaii Telescope, operated by the CNRS of France, CNRC in Canada and the University of Hawaii.

Multi-object spectroscopy down to $I_{AB} = 22.5$ is being obtained over the four fields (“VVDS Wide”), with a deeper area of 1.5 deg^2 in the VVDS-02h and in the Chandra Deep Field South (VVDS-CDFS) covered to $I_{AB} = 24$ (“VVDS Deep”). The first epoch VVDS data consist of more than 11 000 spectra obtained in the VVDS-Deep fields (Le Fèvre et al. 2005a).

One of the key science goals of the VVDS is to measure the evolution of galaxy clustering from the present epoch up to $z \sim 2$. The simplest statistic used for this analysis is the spatial two-point correlation function $\xi(r)$ and its variants, (e.g. Peebles 1980), i.e. the second moment of the galaxy distribution. Given the geometry and selection function of galaxy surveys, however, the practical estimation of $\xi(r)$ from the actual data is not straightforward. Edge effects, sampling inhomogeneities and selection effects all introduce different biases that hamper the survey’s ability to estimate the true underlying clustering process. Moreover, intrinsic systematic uncertainties due to the limited size of the volume of the Universe explored (“cosmic variance”) need to be accounted for when computing realistic error bars on the measured correlation values.

The aim of this paper is to present a comprehensive description of the biases specific to the VVDS, along with the methods we developed to correct for them. The strategy we adopt relies on the construction of realistic “pre-observation” mock catalogs using the *MoMaF* software (Blaizot et al. 2005) and the *GalICS* hybrid model for galaxy formation (Hatton et al. 2003). We then observe these mock catalogs, by mimicking the relevant observational selections and biases. Comparing original and observed mock surveys allows us to (i) quantitatively understand the impact of the different biases inherent to the VVDS data on clustering estimates, and (ii) to explore and validate methods that allow us to recover the original signal. This strategy is possible because *GalICS* predictions have been shown to agree fairly well with a wide range of observations (e.g. Hatton et al. 2003; Blaizot et al. 2004), and is thus expected to yield catalogs realistic enough to carry out a convincing consistency check. Because our mock catalogs contain realistic clustering properties, we can also use them to predict the cosmic variance amplitude in order to compute realistic errors on the clustering estimates we will perform on the real data.

The paper is organized as follows. In Sect. 2 we discuss the different kind of biases expected in the current VVDS first-epoch data. In Sect. 3 we discuss the construction of mock VVDS catalogs from the *GalICS/MoMaF* simulations which assume a flat Cold Dark Matter model with $\Omega_m = 0.333$, $\Omega_\Lambda = 0.667$ and $h = 0.667$. In Sect. 4 we present the definitions of the two-point correlation functions. Then, in Sect. 5 we discuss the details of the error measurement strategy when applied to VVDS. In Sect. 6 we show how the measured two-point correlation function is affected by the features particular to our survey and we discuss the methods developed to correct for these biases and properly estimate the correlation function $\xi(r_p, \pi)$, its projection $w_p(r_p)$, and the correlation length r_0 and slope γ , as a function of redshift. Section 7 summarizes our results.

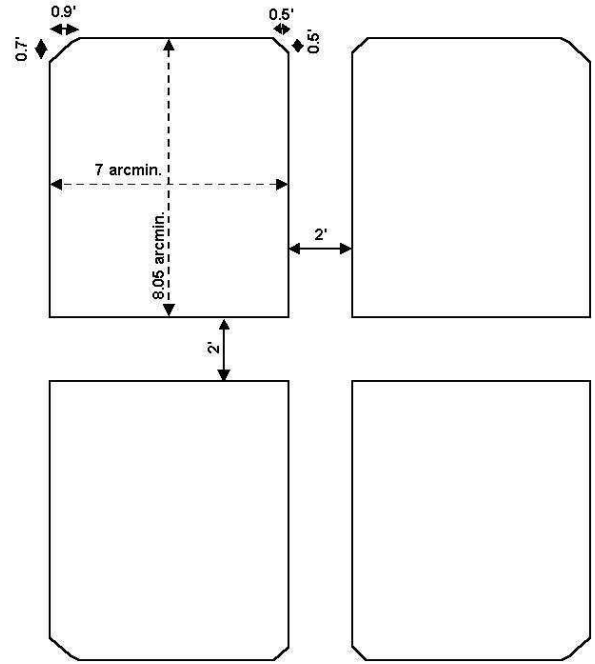


Fig. 1. Lay-out of the VIMOS field of view. INVAR masks with laser-cut slits are placed on the focal plane within the four rectangular areas (“VIMOS channels”).

2. The selection function of VVDS first epoch observations

The first epoch spectra of the VVDS-Deep collected during the 2002 and 2003 campaigns are concentrated within the 02h deep field, and the CDFS (Le Fèvre et al. 2005a). First epoch spectra have been collected for galaxies down to $I_{AB} \leq 24$ in the 0.61 sq degree sub-area of the VVDS-02h field and a region of $21 \times 21.6 \text{ sq arcmin}$ centered on the Chandra Deep Field South (CDFS, Giacomini et al. 2002). The VVDS First Epoch data geometrical lay-out, sampling rate and incompleteness are used as a reference benchmark in this paper.

2.1. Catalog structure and biases

A number of factors, both in the parent photometric catalog from which the target galaxies are selected and in the way the spectroscopic observations are carried out, contribute to create selection effects that bias any estimate of galaxy clustering if not properly accounted for.

1. **Photometric defects.** Some areas are excised from the *I*-band CCD images during their photometric analysis, due to the presence of bright stars or other instrumental effects (e.g. stray-light from a bright star outside the field of view). The resulting photometric galaxy catalog, therefore, features some artificially empty regions.
2. **VIMOS lay-out.** The field of view of the VIMOS spectrograph consists of four $7'$ by $8'$ quadrants, separated by $2'$ gaps, as shown schematically in Fig. 1. At the typical resolution used in the VVDS, between 110 and 150 spectra are collected in each quadrant during a single observation. Clearly, no galaxies are observed over the area

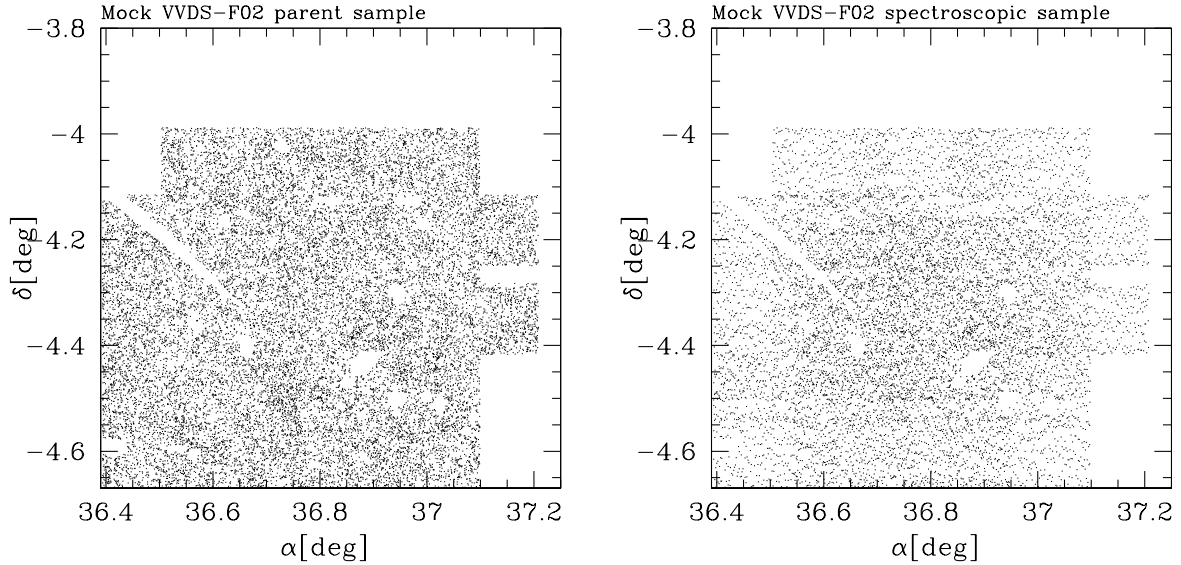


Fig. 2. Galaxy distribution in a mock VVDS-02h catalog, constructed using the *GalICS* simulations with the same lay-out as the 20 observed pointings in the actual first-epoch VVDS field and applying the full range of selection effects present in the data, as e.g. the photometric mask. The left panel shows the parent photometric field, including all objects with $I_{AB} \leq 24$ within the current VVDS-02h boundaries and mask. In the right panel only the objects selected for spectroscopy are shown. Note the density gradient towards the central part of the field, due to multiple passes over the same area.

of the “cross” between the four quadrants, unless one observes the area with a new pointing, shifted with respect to the first one (see below).

3. **Missing quadrants.** For a few pointings, one or two quadrants can be “blind”, i.e. with no spectra observed due to a misplacement of the multi-slit masks during the observations.
4. **Incomplete coverage.** The planned final area is being covered through a mosaic of adjacent pointings. Thus, at any intermediate stage the available spectral data set is distributed in a non-uniform fashion on the sky. The largest contiguous area currently covered in the 02h deep field corresponds to about 0.5 square degrees, with the geometry shown in Fig. 2.
5. **Varying sampling density.** The VVDS observational strategy involves multiple passes over the same area to increase the spectral sampling rate. While a central region of the 02h deep field is exposed 4 times (i.e. it is visited by four independent pointings with different slit masks), the external areas are covered only twice due to the tiling strategy. During subsequent observing runs, the VIMOS pointings are shifted with respect to the previous ones usually by around $2'$, to ensure that the cross visible in Fig. 1 is filled. As a consequence, the mean surface density of observed objects varies across the field.
6. **Optimization of the number of slits and mechanical constraints.** A specific source of bias in the VIMOS observations is introduced by VMMPS – the VIMOS Mask Manufacturing Preparation Software, and specifically by the Super-SPOC code (Bottini et al. 2005). The width of a slit is set to 1 arcsec (or about 5 detector pixels), and its typical length is ~ 6 – 10 arcsec to include both the object of interest and enough information on the sky spectral background to correct for it. The VMMPS software

automatically allocates slits to objects in the input catalog with the goal of maximizing the total number of spectra. In general, this means that the spectroscopic sample is not a random sparse sampling of the clustering pattern over the sky, but a more homogeneous sub-sample. Specifically, VMMPS tends to place objects in rows, so to maximize the number of spectra across the CCD (see Fig. 3), with an additional slight preference towards objects of small angular size. As typical with multi-object spectrographs, the minimum slit size implies that, after one single spectroscopic pass, there is a bias against observing very close angular pairs on the sky. Having multiple passes, however, significantly improves the situation, allowing for very close pairs to be observed in subsequent exposures.

The final spectroscopic sample is thus affected to different degrees by all these factors. Figure 2 shows the current lay-out of the observed pointings in the 02h field, compared to the parent photometric sample over the same area. Features from the two main effects are obvious from Fig. 2: holes in the parent catalog and the varying sampling density in the spectroscopic data, due to the multiple passes over the central area. The “striping” effect due to the slit-placing software is not obvious at this resolution and is better appreciated in Fig. 3, where only one quadrant is displayed.

3. Constructing mock VVDS surveys

The only way to understand the relative importance of the selection biases discussed above and test possible correction schemes is to create and analyze realistic simulations of our survey. Provided these simulations are realistic enough, they allow us (1) to understand quantitatively the magnitude of observational biases on the final statistical quantities to be measured, and (2) to estimate realistic errors that include cosmic

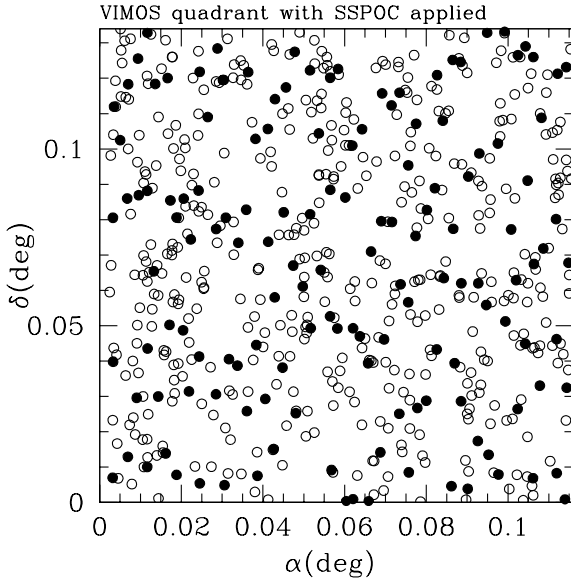


Fig. 3. Spectroscopic targets (filled circles) selected in one of the four VIMOS quadrants from a complete VVDS mock photometric sample (open circles). Note how the optimization software tends to select spectroscopic targets aligned along horizontal rows, while, clearly, very close pairs are not observed. Typically, however, 4 independent observations are conducted on the same area, each with a similar target layout, but shifted by a few arcminutes. This significantly reduces both the alignment and proximity effects. The residual bias is then further corrected by the weighting scheme discussed in Sect. 4. Overall, the four passes produce a typical sampling rate of one galaxy in four.

variance. Both these points require that mock observations contain a spatial distribution of galaxies consistent with VVDS observations – so as to measure clustering and cosmic variance – along with realistic photometric and physical properties of simulated galaxies – so as to mimic selection effects. The *GalICS* model for galaxy formation (Hatton et al. 2003) along with the *MoMaF* mock observing tool (Blaizot et al. 2005) fulfill these requirements and we thus use them to build “pre-observation” catalogs that we then “observe” by progressively adding all the VVDS observational biases and selections.

In this section, we first describe the *GalICS* simulation that we use, before discussing how we build simulated VVDS observations that account for all identified biases.

3.1. The *GalICS* simulations

GalICS (for *Galaxies In Cosmological Simulations*, see Hatton et al. 2003) is a model of hierarchical galaxy formation which combines high resolution cosmological simulations to describe the dark matter content of the Universe with semi-analytic prescriptions to deal with the baryonic matter.

The cosmological N -body simulation we refer to throughout this paper assumes a flat cold dark matter model with a cosmological constant ($\Omega_m = 0.333$, $\Omega_\Lambda = 0.667$). The simulated volume is a cube of side $L_{\text{box}} = 100 h^{-1}$ Mpc, with $h = 0.667$, containing 256^3 particles of mass $8.272 \times 10^9 M_\odot$, with a smoothing length of 29.29 kpc. The power spectrum was set in agreement with the present-day abundance of rich

clusters ($\sigma_8 = 0.88$, from Eke et al. 1996), and the DM density field was evolved from $z = 35.59$ to $z = 0$, outputting 100 snapshots spaced logarithmically in the expansion factor.

GalICS builds galaxies from this simulation in two steps. First, halos of DM containing more than 20 particles are identified in each snapshot using a friend-of-friend algorithm. Their merging history trees are then computed following the constituent particles from one output to the next. Second, baryons are evolved within these halo merging history trees according to a set of semi-analytic prescriptions that aim to account for e.g. heating and cooling of the gas within halos, star formation and its feedback on the environment, stellar population evolution and metal enrichment, formation of spheroids through galaxy mergers or disc instabilities.

Three main points make *GalICS* particularly suitable for this study. First, this model yields a wide range of predictions, among which luminosities (in many bands from the UV to the sub-mm), physical properties (such as sizes of galaxies), and the positions of galaxies within the simulation snapshots. Second, these properties have been shown to be in a rather good agreement with various observations (e.g. Hatton et al. 2003; Blaizot et al. 2004). Third, mock observations are readily available from the *GalICS* Project’s web-page¹. These mock observations include 50 catalogs of 1×1 sq deg that contain all the information we need in this study: apparent magnitudes in the *BVRI* filters used at the CFHT, apparent sizes of the galaxies, angular coordinates in the mock sky, and redshifts.

Before using *GalICS* mock samples, it is useful to state their limitations (see however Blaizot et al. 2005, for a thorough description of these). There are mainly three shortcomings to mock catalogs made using *GalICS*. First, because of the finite mass resolution of the root simulation, faint galaxies are not well described, or even missed when they lie in unresolved haloes. This is not an issue for the present study, however, because the VVDS detection limit is brighter than *GalICS*’s resolution. Second, because mock catalogs are built from a simulation of a finite volume, they involve replications of this volume, along and perpendicular to the line of sight. These replications lead to some negative bias in the correlation functions, of at most $\sim 10\%$. This is not a concern in this paper, because we just need an approximate match with the observed data in order to perform an *internal* consistency check. *GalICS* catalogs do provide an adequate match. Third, the mock catalogs do not describe density fluctuations on scales larger than the size of the simulated volume ($\sim 100 h^{-1}$ Mpc). This implies that cosmic variance estimates are likely to be under-estimated if the volume probed by a mock catalog is larger than the simulated volume. This under-estimate, however, depends on the galaxy population considered: it will be large for rare objects and small for “normal” galaxies. In other words, because cosmic variance is basically given by the integral of the correlation function over the survey, the error on the estimated cosmic variance depends on how much of this integral we miss, that is, on how strongly the studied galaxies are clustered. From Fig. 9, it can be seen that the size of the simulation is enough for this under-estimate to be small at the scales

¹ <http://galics.iap.fr>

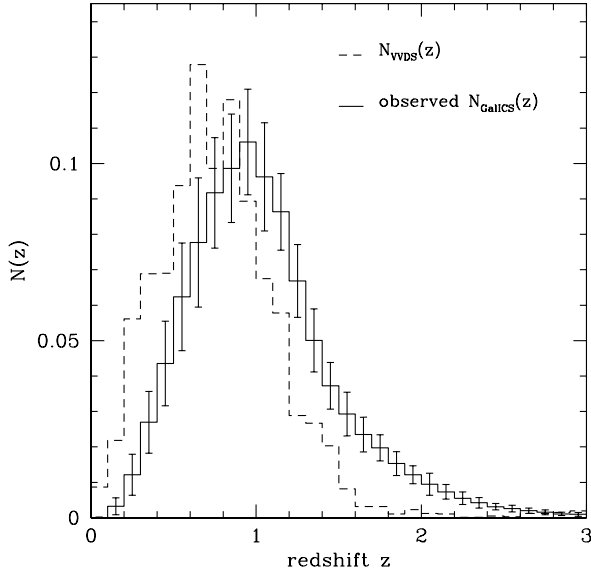


Fig. 4. Average redshift distribution in the 50 mock VVDS-02h surveys, normalized by the number of objects in each cone, compared to the redshift distribution of the observed VVDS galaxies. Note how the semi-analytic model of galaxy formation used to construct the *GalICS* simulations differs from the real data. This is not a concern for the purposes of this work: first, we are performing *internal* tests of the effect of observing biases and on their correction, which depends on the small-medium scale clustering properties. Second, when error bars are estimated for a specific redshift slice, their amplitude is re-normalized accordingly, to account for the different number of galaxies.

we consider (i.e. from 0.1 to $10 h^{-1}$ Mpc). The dispersion found among the 50 *GalICS* cones is thus expected to be a good estimate of cosmic variance. The mean number of galaxies with $17.5 < I_{AB} < 24$ in the artificial catalogs is 77 396. The average redshift distribution of these 50 cones is shown in Fig. 4, along with the VVDS first epoch $N(z)$ (Le Fèvre et al. 2005a).

We note that the redshift distribution of the simulated galaxies differs significantly from that observed by the VVDS for the real Universe. This is simply telling us that the semi-analytic galaxy formation model adopted to construct the *GalICS* simulations, while adequately reproducing a number of observed features (see Blaizot et al. 2005) is not 100% correct. This, however, is of no importance for the current analysis, as our main goal is to test the internal differences in the measured quantities when either the original parent sample or the final spectroscopic sample are observed. The accuracy of these tests depends essentially on the small-scale properties of the simulated galaxies (like the mean inter-galaxy separation and clustering), rather than on the global redshift distribution. Conversely, in the estimate of error bars the difference in absolute numbers between the real and simulated samples within a given redshift slice will clearly have to be taken into account.

3.2. CCD photometric mask

Bright (often saturated) stars represent a practical obstacle to accurate galaxy photometry and their diffused light can affect large areas of a CCD astronomical image. All such areas were

excised from the VVDS photometric catalogs: there are no sources in these regions (McCracken et al. 2003). Similarly, a “dead” area in the 02h field has been produced by a beam of scattered light that crosses a large part of the field from North–East to South–West. In total, a few percent of the total area are lost due to these defaults. The information on these “holes” in the photometric catalog is stored in a FITS binary mask, with null values corresponding to dead pixels. We have used this mask on the mock samples to exactly reproduce the pattern of the observed data in our simulations.

3.3. Effect of galaxy angular sizes

In order to maximize the number of spectroscopic targets, the Super-SPOC software (Bottini et al. 2005) makes a choice of a targeted galaxy based also on the galaxy projected angular radius along the slit direction. This means that smaller galaxies are sometimes preferred as they allow the program to increase the number of targets. Any realistically simulated spectroscopic sample must take this into account. Therefore, we have computed for each simulated galaxy in *GalICS* a realistic angular radius, using the following procedure.

GalICS describes galaxies with three components: a disc, a bulge and possibly a nuclear starburst. For each of these, the model predicts the mass and a scale-length that assumes the disc is exponential while the other two spheroidal components follow a Hernquist profile (Hernquist 1990). We used these sizes to define an overall radius for each galaxy, which encloses 90% of the total mass. Assuming that light has the same distribution as mass, we then convert this radius to an apparent angular size, assuming the above-mentioned cosmology.

3.4. Artificial stars

The VVDS spectroscopic targets are selected purely on magnitude, $I_{AB} \leq 24$ and $I_{AB} \leq 22.5$ in the Deep and Wide parts of the survey, respectively, without any a priori star-galaxy separation. This avoids biases against compact galaxies and AGNs which may be introduced at faint magnitudes by unreliable star-galaxy classification based on morphology. Consequently, our spectroscopic sample is contaminated by stars. About 8.5% of the collected spectra in the VVDS-Deep are stars and are discarded (the exact number depending on galactic latitude can be as high as 20% in some cases for the “Wide” survey). These stars obviously have no impact on the clustering analysis. Their only effect is to reduce the total number of targeted galaxies, thus slightly affecting the overall statistics by increasing the expected variance. Since our aim here is to precisely quantify the biases and uncertainties on galaxy correlations computed from the final spectroscopic sample, and compare them to the original parent sample, we decided to also take into account this small contribution. We therefore added to the artificial survey fields a set of simulated stars.

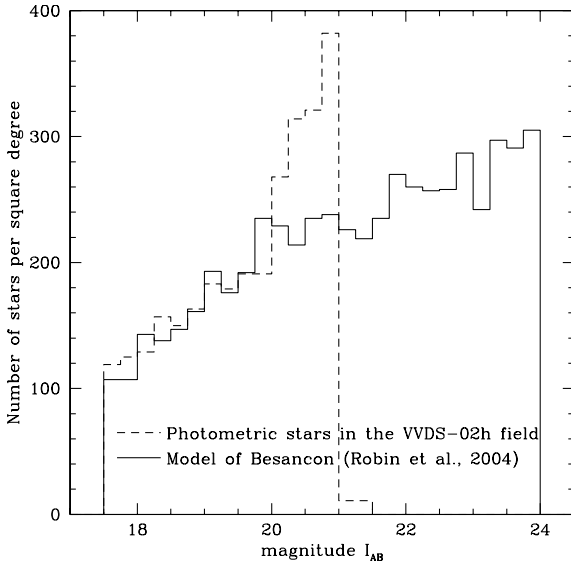


Fig. 5. Number counts of artificial stars added to the *GalICS* simulation, compared to the actual counts of stars in the VVDS-02h field, identified morphologically from the photometric data. The excess in the VVDS above $I_{AB} = 20$ is due to the inability of the morphological compactness criteria to discriminate stars from galaxies and QSOs at faint magnitudes. When this is taken into account, the models from Robin et al. (2003) reproduce very well the actual distribution of stellar objects in the VVDS.

Using the on-line tool of Robin et al. (2003)² we created a one-square-degree catalog of artificial stars with $17.5 \leq I_{AB} \leq 24$, which was added to the artificial galaxy photometric catalogs. Figure 5 shows the number counts of the added stars, compared to the observed distribution at bright magnitudes in the 02h field (as identified by *S-extractor*, Bertin & Arnouts 1996). The observed excess above $I_{AB} = 20$ in the 02h field is the effect of mis-classified galaxies and QSOs, which also corroborates our choice of excluding any pre-selection for the VVDS spectroscopy, to avoid throwing these objects away.

As this parameter is used by VMMPs, apparent angular radii have also been assigned to artificial stars, using the observed distribution of stellar sizes in the 02h field, identified photometrically down to $I_{AB} = 21$ and spectroscopically at fainter magnitudes. This range of apparent stellar radii corresponds to the sizes of the point spread function (“seeing”) at the faint Kron radii measured for stars by *S-extractor*.

3.5. Spectroscopic success rate

Objects selected by the slit-positioning code do not yet form the final redshift catalog. For some of the objects, redshift measurements are impossible, usually because of poor signal-to-noise. This incompleteness is clearly a function of magnitude. We define the spectroscopic success rate as the ratio of the number

² The *Model of stellar population synthesis of the Galaxy* developed by Robin et al. (2003) produces a reliable catalog of stars with appropriate number counts and magnitudes in the visible and near-infrared spectral ranges in the Johnson-Cousins and Koornneef systems, respectively.

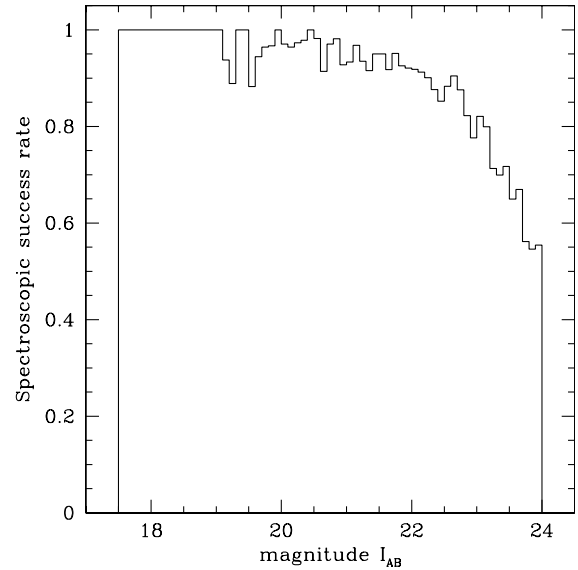


Fig. 6. Spectroscopic success rate per magnitude bin in the VVDS 02h field, including only those redshifts used for the clustering analysis.

of redshifts used for clustering analysis to the total number of spectroscopically observed objects. Figure 6 shows the spectroscopic success rate as a function of magnitude, which corresponds in practice to the probability of measuring the correct redshift of a galaxy as a function of its magnitude in the current observational configuration. Overall, this shows that we are able to obtain a redshift for more than 80% of the targeted objects between $I_{AB} = 17.5$ and 24. We therefore apply this same probability function to our mock “observed” catalogs, rejecting the corresponding fraction of targeted objects. We make the simplifying assumption that the spectroscopic success rate is the same for all galaxy types.

3.6. VIMOS spectral resolution

The last point to be taken into account to produce a fully realistic mock redshift catalog is the resolution of the VIMOS spectrograph in the set-up used for the VVDS (Low-resolution RED Grism, $R \simeq 230$) which translates into a typical rms error on the measured redshift which is around $\sigma_{cz} \simeq 275 \text{ km s}^{-1}$. We therefore added to the final set of mock redshifts a Gaussian-distributed dispersion with the same rms and zero mean.

3.7. Overall properties of mock VVDS surveys

All of the steps described above have been applied to each of the 50 one-square-degree *GalICS* surveys, producing a corresponding number of mock redshift samples which reproduce with fidelity the lay-out, properties and biases of the first-epoch VVDS 02h sample.

Figure 7 shows that, despite the slight bias of SSPOC towards choosing smaller (and therefore fainter) objects, the redshift distribution $N(z)$ of the final spectroscopic samples is unbiased with respect to the original complete *GalICS* one-square-degree survey. The difference observed in Fig. 4 between the original and observed simulated cones is therefore

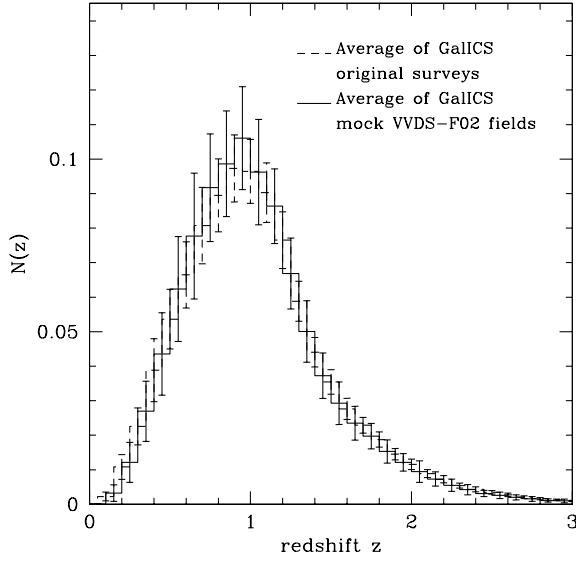


Fig. 7. Average redshift distribution in the *GalICS* mock catalogs before and after the full observing strategy is applied. No bias in the redshift distribution is observed.

only the result of the model of galaxy formation adopted for the simulation, and not of a selection effect. There was no way we could introduce, e.g., a stronger incompleteness in the final $N(z)$ at $z > 1$.

4. Two-point correlation statistics

4.1. General estimator

The two-point correlation function $\xi(r)$ is defined as the excess probability above random that a pair of galaxies is observed at a given spatial separation r (Peebles 1980).

It is the simplest statistical measurement of clustering, as a function of scale, and it corresponds to the second moment of the distribution. Various recipes have been proposed to estimate two-point correlation functions from galaxy surveys, in particular to minimize the biases introduced by the finite sample volume, edge effects, and photometric masks (Hamilton 1993; Landy & Szalay 1993). Here we adopt the Landy-Szalay estimator, that expresses $\xi(r)$ as

$$\xi(r) = \frac{N_R(N_R - 1)}{N_G(N_G - 1)} \frac{GG(r)}{RR(r)} - 2 \frac{N_R - 1}{N_G} \frac{GR(r)}{RR(r)} + 1. \quad (1)$$

In this expression, N_G and N_R are the mean density (or, equivalently, the total number) of objects respectively in the galaxy sample and in a catalog of random points distributed within the same survey volume and with the same redshift distribution and angular selection biases; $GG(r)$ is the number of independent galaxy-galaxy pairs with separation between r and $r+dr$; $RR(r)$ is the number of independent random-random pairs within the same interval of separations and $GR(r)$ represents the number of galaxy-random pairs.

4.2. Redshift-space correlations

We know that the three-dimensional galaxy distribution recovered from a redshift survey is distorted due to the effect of

peculiar velocities. For this reason, the redshift-space separation s differs from the true physical comoving separation r between two galaxies. Since random velocities affect only redshift and not position on the sky, the stretching occurs only radially. Redshift distortions can be measured and separated from true spatial correlations by computing the function $\xi(r_p, \pi)$, where the separation vector of a pair of galaxies s is split into two components: π and r_p , respectively parallel and perpendicular to the line of sight. Given two objects at redshifts z_1 and z_2 , with observed radial velocities $v_1 = cz_1$ and $v_2 = cz_2$ (c being the speed of light), we can define (Fisher et al. 1994) the line of sight vector $\mathbf{l} \equiv (\mathbf{v}_1 + \mathbf{v}_2)/2$ and the redshift difference vector $\mathbf{s} = \mathbf{v}_1 - \mathbf{v}_2$, and also:

$$\pi \equiv \frac{\mathbf{s} \cdot \mathbf{l}}{H_0 |\mathbf{l}|}, \quad r_p^2 \equiv \frac{\mathbf{s} \cdot \mathbf{s}}{H_0^2} - \pi^2. \quad (2)$$

Equation (1) can be generalized to the case of $\xi(r_p, \pi)$, if we count the number of pairs in a grid of bins Δr_p and $\Delta \pi$ instead of singular bins Δr or Δs .

Observed distortions in galaxy surveys can be separated into two main contributions: on small scales, the distortion is dominated by random internal velocities in groups and clusters, causing a stretching of $\xi(r_p, \pi)$ along the π direction (the so-called “fingers of God” effect). On large scales, on the other hand, $\xi(r_p, \pi)$ contours tend to be flatter, due to coherent infall of galaxies onto large-scale overdensities, known as the “Kaiser effect” (Kaiser 1987). The latter is a weak effect and needs very large samples to be seen with sufficient accuracy, as shown by the 2dF survey (Hawkins et al. 2003).

4.3. Projected correlation function $w_p(r_p)$

We can recover the real-space correlation function $\xi(r)$ by projecting $\xi(r_p, \pi)$ along the line of sight, onto the r_p axis. In this way we integrate out the dilution produced by the redshift-space distortion field and obtain a quantity, $w_p(r_p)$, which is independent of the redshift-space distortions:

$$w_p(r_p) \equiv 2 \int_0^\infty \xi(r_p, \pi) d\pi = 2 \int_0^\infty \xi_\circ \left[(r_p^2 + y^2)^{1/2} \right] dy. \quad (3)$$

In the right-hand side of the equation, ξ_\circ is simply the usual real-space two-point correlation function $\xi(r)$, evaluated at the specific separation $r = \sqrt{r_p^2 + y^2}$. If we now assume a power-law model

$$\xi(r) = \left(\frac{r}{r_0} \right)^{-\gamma}, \quad (4)$$

with γ being the slope of the correlation function and r_0 the correlation length, the integral can be computed analytically, giving as a result

$$w_p(r_p) = r_p \left(\frac{r_0}{r_p} \right)^\gamma \frac{\Gamma\left(\frac{1}{2}\right) \Gamma\left(\frac{\gamma-1}{2}\right)}{\Gamma\left(\frac{\gamma}{2}\right)}, \quad (5)$$

where Γ is Euler’s Gamma function.

5. Error estimate and fitting technique

5.1. Error bars on correlation functions

Ideally, if the studied data set consisted of a large enough number of statistically independent pairs, such that the central limit theorem applies, then the distribution of estimates of ξ in an ensemble of similar samples should be Gaussian. The 1σ uncertainty – the “cosmic error” – in ξ would then be the square root of its variance $\langle \Delta\xi^2 \rangle$ (Peebles 1973). However, the theoretical expression for $\langle \Delta\xi^2 \rangle$ depends on the poorly known and difficult to measure four-point correlation function. Moreover, since the measured ξ is not exactly coincident with the theoretical ξ , we expect its uncertainty to be also somewhat different from the value provided by the theory. This effect is known as a cosmic bias.

A few different ways of estimating errors on two-point correlation functions have been used in the literature (for a wider discussion, see e.g. Hamilton 1993; Fisher et al. 1994; Bernardeau et al. 2002). The case closest to the ideal situation is when the survey is large enough that it can be split into a number of sub-samples. Correlations are then estimated independently for each of these, and error bars for the parent sample computed as the rms values. This has been for example the case of the angular correlation function from the APM survey (e.g. Maddox et al. 1990). However, the number of sub-samples cannot be large, otherwise the explored scales will be significantly reduced with respect to the parent survey. The consequence is that the variance is typically overestimated and these represent usually upper limits to the true errors.

Simple Poissonian errors (e.g. proportional to the square root of the total number of galaxy pairs in each bin) underestimate the error bars substantially. Statistical corrections were proposed (Kaiser 1986) by multiplying Poissonian errors by a factor $1 + 4\pi n J_3$, with n being the number density of objects and $J_3 = \int_0^{r_j} r^2 \xi(r) dr$, where we assume that the actual correlation function vanishes for $r \geq r_j$. However, this method also tends to give relatively small errors (Fisher et al. 1994).

Over the last twenty years a widely used method has been the so-called “bootstrap resampling” (Barrow et al. 1984). It is based on the idea of “perturbing” the data set, by randomly creating a large number of comparable “pseudo data-sets”, which differ only slightly from the original sample. If this contains N objects, then each bootstrap sample is created selecting N of these, but allowing for multiple selections of the same object. This means that some objects will not be included in one given pseudo data-set, while others will be counted twice or three times. This is a good test of the robustness of measured correlations, especially on large scales where having a large number of pairs does not always mean a robust measurement: consider for example the case of a single isolated galaxy at a separation of \bar{r} from a cluster containing 1000 galaxies. $\xi(\bar{r})$ will contain a large number of pairs, however only one will be independent. On the other hand, bootstrap errors often tend to over-estimate the theoretical variance $\langle \Delta\xi^2 \rangle$. In general, however, despite debates on their theoretical justification, they have represented a practical way to obtain error bars in correlation analysis which are not far from the true ones.

The use of bootstrapping became less and less popular in recent years, with the advent of large N -body simulations, reproducing the matter distribution over significant volumes of the Universe. Coupled to physically sound definitions of “galaxies”, these allowed the construction of sets of independent mock surveys, from which ensemble errors could be computed from the scatter in the different catalogs. This is the same technique used to construct our VVDS mock surveys. Clearly, a good match is necessary between the volume and resolution of the simulation, on one side, and the depth and size of the survey on the other. Furthermore, the power spectrum of the simulation must provide a realistic description of long waves, so to properly include cosmic variance. Progress both in our knowledge of structure on the largest scales and in the size and resolution of N -body simulations has improved on early applications of this technique (Fisher et al. 1994). For this reason, since the *GallCS* simulations are available, we could use this as our main method for error estimation.

However, as we detail below, the covariance matrix reconstructed from the simulations cannot be applied in a straightforward way to the observed data. Indeed, our fitting technique, discussed below, handles the covariance matrix to properly account for bin-to-bin correlations when fitting correlation functions: when the covariance matrix extracted from the set of 50 mock VVDS surveys is used (after proper normalization of the average values), the fit is often unstable. In other words, the covariance matrix produced by the ensemble of mock surveys, although providing sufficiently realistic diagonal elements, has off-diagonal non-zero values which differ from those pertaining to the data sample (which of course are unknown). For this reason, we modified our strategy and resort to the bootstrap technique to *estimate the bin-to-bin covariance*. This means that our error bars on the estimated correlation functions are obtained via the more reliable scatter between the mock surveys, but a bootstrap is used to estimate the off-diagonal terms of the covariance matrix.

5.2. Fitting correlation functions

It is well known that fitting of correlation functions like $\xi(s)$ or $w_p(r_p)$ cannot be performed via the standard least-squared method, due to the correlation existing among the different bins. The procedure we adopted to estimate the power-law parameters of $\xi(r)$, r_0 and γ from the projected function $w_p(r_p)$, using Eq. (5) follows Fisher et al. (1994) and Guzzo et al. (1997), with some specific differences that are described in the following.

Let us consider a given redshift slice $[z_1 - z_2]$. Within this same interval, we estimate the correlation function $\xi(r_p, \pi)$ from: 1) 50 mock VVDS surveys; 2) the real VVDS data; 3) N_{boot} (typically 100) bootstrap resamplings of the VVDS data. We then compute, for each of these estimates, $w_p(r_p)$, projecting $\xi(r_p, \pi)$ along the line of sight (Eq. (3)), with an upper integration limit π_{max} , chosen in practice so that it is large enough to produce a stable estimate of w_p . Similarly to other authors (see e.g. Guzzo et al. 1997), we find $w_p(r_p)$ quite insensitive to the choice of π_{max} in the range of

$15 h^{-1} \text{Mpc} < \pi_{\text{max}} < 25 h^{-1} \text{Mpc}$ for $r_p < 10 h^{-1} \text{Mpc}$. Too small a value for this limit would miss small-scale power, while too large a value has the effect of adding noise into w_p . After a set of experiments we have chosen $\pi_{\text{max}} = 20 h^{-1} \text{Mpc}$.

In the following, we call $w_p^k(r_i)$ the value of w_p , computed at $r_p = r_i$ in the cone k , where $1 \leq k \leq N_{\text{GalICS}} = 50$ if we consider the *GalICS* data or $1 \leq k \leq N_{\text{boot}}$ if we consider the bootstrap data. If not otherwise mentioned, $N_{\text{boot}} = 100$ is used.

Whether we consider the mock or bootstrap samples, we can always compute the associated covariance matrix, \mathbf{C} , between the values of w_p in i th and k th bins:

$$\mathbf{C}_{ik} = \left\langle \left(w_p^i(r_i) - \langle w_p^i(r_i) \rangle_j \right) \left(w_p^k(r_k) - \langle w_p^k(r_k) \rangle_j \right) \right\rangle_j, \quad (6)$$

where “ $\langle \rangle_j$ ” indicates an average over all bootstrap or mock realizations. When the correlation function is computed from a finite sample, the values of $\xi(r)$ (or $w_p(r)$) at different separations are not independent³ For this reason one cannot use a straightforward χ^2 minimization – which assumes that all bins are independent and that the errors follow the Gaussian distribution – to find the best-fit parameters of a model to the observed data. However, \mathbf{C} is symmetric and real and therefore can be diagonalized by a unitary transformation if its determinant is non-vanishing. In practice, the estimated functions are oversampled, \mathbf{C} is not singular and therefore can be inverted by a simple Cholesky decomposition (Numerical Recipes, Press et al. 1992, Volume 1, Chapter 2)⁴. Then, if we now call $\mathbf{H} = \mathbf{C}^{-1}$, we can fit w_p^{VVDS} by minimizing a generalized χ^2 , which is defined as

$$\chi^2 = \sum_{i=1}^{N_D} \sum_{j=1}^{N_D} \left(w_p^{\text{mod}}(r_i) - w_p^{\text{VVDS}}(r_i) \right) H_{ij} \left(w_p^{\text{mod}}(r_j) - w_p^{\text{VVDS}}(r_j) \right), \quad (7)$$

as a function of the two free parameters r_0 and γ of $w_p^{\text{mod}}(r_p)$.

In principle, the complete process could be done using only our set of 50 mock VVDS surveys. In practice, as explained above, the bin-to-bin covariance obtained from the *GalICS* mock samples does not provide a statistically stable matrix to be used with the generalized χ^2 method. Therefore, we most appropriately used the covariance matrix obtained from the N_{boot} bootstrap resamplings of the galaxy data set.

This provides the best solution for $(r_0, \gamma)_{\text{data}}$ that minimizes the error contour $\chi_{\text{boot}}^2(r_p, \gamma)$. At the same time, however, we use 50 mock surveys to obtain the most realistic error contours $\chi^2(r_p, \gamma)$ on our estimated $(r_0, \gamma)_{\text{data}}$, as these – unlike bootstrap errors – include cosmic variance.

The final error contours, therefore, are obtained fitting the mean of the 50 w_p mock VVDS surveys, using a covariance

matrix computed from the same 50 w_p . This process provides a solution for $(r_0, \gamma)_{\text{GalICS}}$ associated with the error contours $\chi_{\text{GalICS}}^2(r_p, \gamma)$. We then re-center these contours around $(r_0, \gamma)_{\text{data}}$ with the renormalization $r_p \leftarrow r_p \times (r_0^{\text{GalICS}}/r_0^{\text{data}})$ and $\gamma \leftarrow \gamma \times (\gamma^{\text{GalICS}}/\gamma^{\text{data}})$.

To take into account the different $N(z)$ of *GalICS* and VVDS, we multiply the error contour χ_{GalICS}^2 computed for each redshift slice by a factor $N_{\text{VVDS}}/N_{\text{GalICS}}$, where N_{VVDS} is the number of VVDS galaxies and N_{GalICS} is the number of *GalICS* galaxies in this redshift slice.

The error bars computed as above for each $w_p(r_i)$ value correspond to the rms of the 50 $w_p^k(r_i)$, normalized to the data.

6. Biasing effects and their removal

We now quantitatively establish the impact of the VVDS selection effects on the measured correlations and the accuracy of our correcting scheme, using the *GalICS* mock samples.

6.1. Impact on angular correlations

As we have seen in the previous section, the biases and selection effects due to the observing strategy and instrumental limitations affect the properties of the angular distribution of objects, with respect to a random sub-sampling of galaxy clustering process. It is therefore the angular correlation function $\omega(\theta)$ that will primarily reflect these biases. Clearly, there is no specific scientific reason to measure the angular correlation function from the spectroscopic sample, as this can be done more easily and with much greater confidence using the full VVDS photometric catalog (McCracken et al. 2003). $\omega(\theta)$ allows us to illustrate the level of distortions introduced by our angular selection function.

To this end, Fig. 8 shows the angular correlation function computed from one mock VVDS redshift survey without correcting for these effects (i.e. using a random sample which simply follows the geometrical borders of the galaxy sample, as one would do for a homogeneous angular selection), compared to that of the original mock catalog. We used the angular version of the Landy-Szalay estimator (Eq. (1)), without taking into account any incompleteness on any scales. The comparison to the parent survey $\omega(\theta)$ reveals the very strong distortions introduced over a wide range of angular scales.

6.2. Correction scheme

The biases discussed so far involve introducing two types of corrections which we discuss in detail in this section.

1) *Global correction.* To account for the effects of uneven boundaries and varying sampling rate we construct a random catalog, which consists of the same number of separately created pointings as the galaxy sample, thus reproducing the global “exposure map” (i.e. number of multiple passes over a given point of the sky) and the corresponding large-scale surface density variations of the galaxy redshift sample. The holes and excised regions in the photometric sample are similarly taken into account by applying the same binary mask to the random sample. These first-order corrections account already

³ For example, imagine that one galaxy is removed from the sample: this galaxy contributes pairs at many different separations, thus affecting virtually all bins in the correlation function.

⁴ Note that if the number of bins we want to fit, i.e. the size of the matrix, were greater or equal to the number of realizations then, even if the matrix remains symmetric, the vectors would not be independent and the matrix \mathbf{C} could not be inverted.

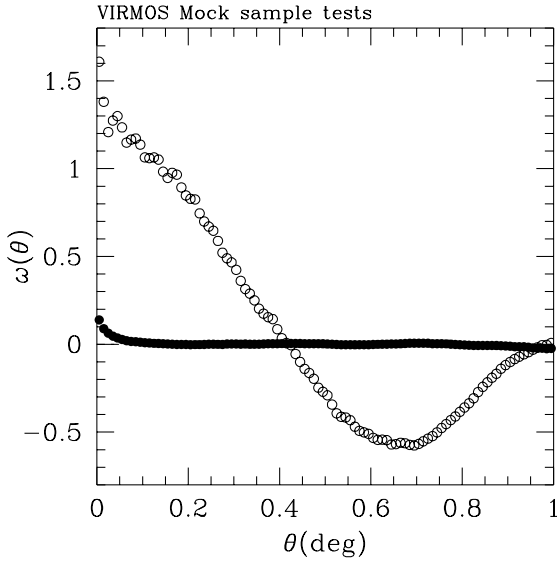


Fig. 8. Impact of the observational process on the estimate of the angular two-point correlation function $\omega(\theta)$ for one mock VVDS survey (open circles), compared to that of the original parent field (filled circles), for one mock VVDS cone. The large distortion, introduced by the observing strategy affects practically all angular scales.

for most of the observational biases. When taken into account, they reduce most of the negative effects of the observing strategy on the correlation functions, in particular the global overestimation of correlation functions (see Figs. 9–11).

2) *Small scale correction.* What remains to be corrected is the slight bias introduced by the slit-positioning software and the mechanical limitations (slit size, closeness of slits and so forth). We have seen that the SSPOC selection is not an entirely random sampling of the actual angular distribution of objects, but rather a more homogeneous sub-set, preferentially concentrated along specific rows. This selection affects primarily the small-scale values of the correlation function, corresponding to the typical slit size: with only one spectroscopic pass, pairs of galaxies with separation smaller than the slit size will always have only one galaxy observed, and thus their contribution to ξ will be lost. With repeated passes this problem is alleviated, as the software chooses each time different objects (except for a small number of objects observed twice for error checking purposes). Using the full 2D information available from the parent photometric catalog (that tells us how many galaxies on the sky have been missed in the spectroscopic sample), we developed a weighting scheme that weighs each targeted galaxy proportionally to its “representativity” in terms of local angular pair density.

Let us therefore consider a circular region of radius θ_w around a galaxy i located within a specific redshift slice k , and define inside θ_w the following quantities:

$n_{\text{gal}}(i)$ – the number of galaxies in the parent photometric catalog;

$n_z(i)$ – the number of galaxies with measured redshift;

$n_{\text{in}}(i)$ – the subset of these belonging to the same redshift slice as the central galaxy;

$n_{\text{exp}}(i)$ – the number of galaxies expected to belong to the same redshift slice, which can be written as

$$n_{\text{exp}}(i) = n_{\text{in}}(i) + n_{\text{rem}}(i), \quad (8)$$

with n_{rem} being the fraction of unobserved neighboring galaxies in the parent photometric catalog expected to belong to the same redshift slice. This number can be written as

$$n_{\text{rem}}(i) = [n_{\text{gal}}(i) - n_z(i)] * P_{\text{slice}}, \quad (9)$$

where P_{slice} is the probability that a generic measured galaxy belongs to that specific redshift slice. Here one can make the reasonable assumption that the observed redshift distribution N_z is sufficiently well sampled as to provide, when averaged over a suitably chosen radius θ_A , an unbiased estimate of P_{slice} for any k th slice as

$$P_{\text{slice}} = \frac{N_{z,k}(<\theta_A)}{N_{z,\text{total}}(<\theta_A)}. \quad (10)$$

The choice of θ_A is clearly critical, as it has to be large enough to allow a proper sampling of existing structures along the line of sight (and thus minimize the noise introduced by the weight), but also small enough not to dilute the effect of single structures within one redshift slice. In practice, given the current size of the 02h field (~ 0.5 square degrees), we have obtained the best results using $\theta_A = 30'$, which encloses virtually the entire field. Note also that $n_{\text{gal}}(i)$, i.e. the number of galaxies in the parent catalog, will be given by $n_{\text{gal}}(i) = n_{\text{all}}(i) * f_{\text{gal}}$, with f_{gal} being the probability that a randomly chosen object from the photometric catalog is not a star but a galaxy and $n_{\text{all}}(i)$ – the number of all locally observed objects in this catalog. For the actual VVDS 02h field, this probability has been estimated to be $f_{\text{gal}} = 0.92$.

The construction of the actual weight to recover the loss of small-scale pairs produced essentially by the proximity bias is not unequivocal. After several experiments with weighting by local densities (of expected vs. observed spectra), we obtained the best results weighting by pairs. The two-point correlation function being a pair-weighted statistic, we constructed our weight $w(i)$ for a given galaxy i from the ratio of the expected to the measured number of pairs within θ_w . Specifically, if one wants the local angular pair density to be conserved, each pair should be counted as:

$$w(i) * w(j) = \frac{n_{\text{exp}}(i) * (n_{\text{exp}}(j) - 1)}{n_{\text{in}}(i) * (n_{\text{in}}(j) - 1)}. \quad (11)$$

And, consequently, a single object is assigned a weight

$$w(i) = \sqrt{\frac{n_{\text{exp}}(i) * (n_{\text{exp}}(i) - 1)}{n_{\text{in}}(i) * (n_{\text{in}}(i) - 1)}}. \quad (12)$$

To define the optimal angular size θ_w defining the “neighborhood” of a galaxy, we experimented with different values in the range $5''$ to $1'$. Not surprisingly, the best correction is obtained for θ_w in the range 30 – $45''$, which is comparable to the length of the VIMOS spectra as projected on the sky. In all computations presented here, we adopted the value $\theta_w = 40''$.

The following sections will present the results of extensive tests of this correction scheme, based on the *GalICS* mock VVDS surveys.

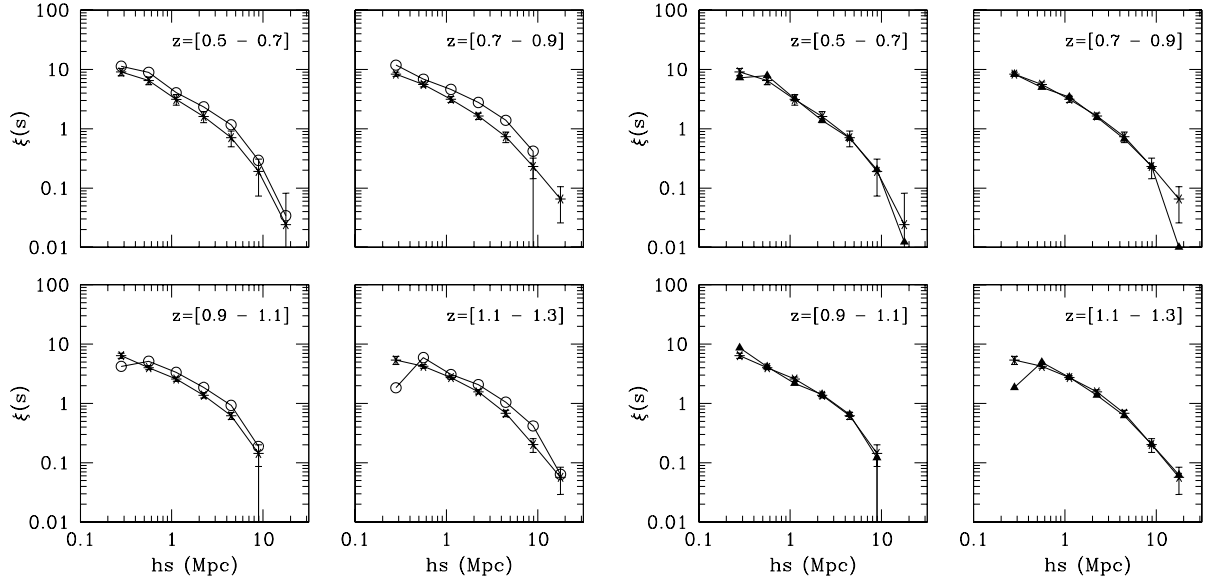


Fig. 9. Redshift-space two-point correlation function $\xi(s)$ for one mock VVDS-02h field, computed in four redshift bins. The true $\xi(s)$ computed for the whole parent sample (stars) is compared to that measured from the “observed” sample, first without any correction (open circles, *left four panels*) and then applying our correction scheme (triangles, *right four panels*). Error bars are the 1σ ensemble rms among the 50 VVDS mock samples.

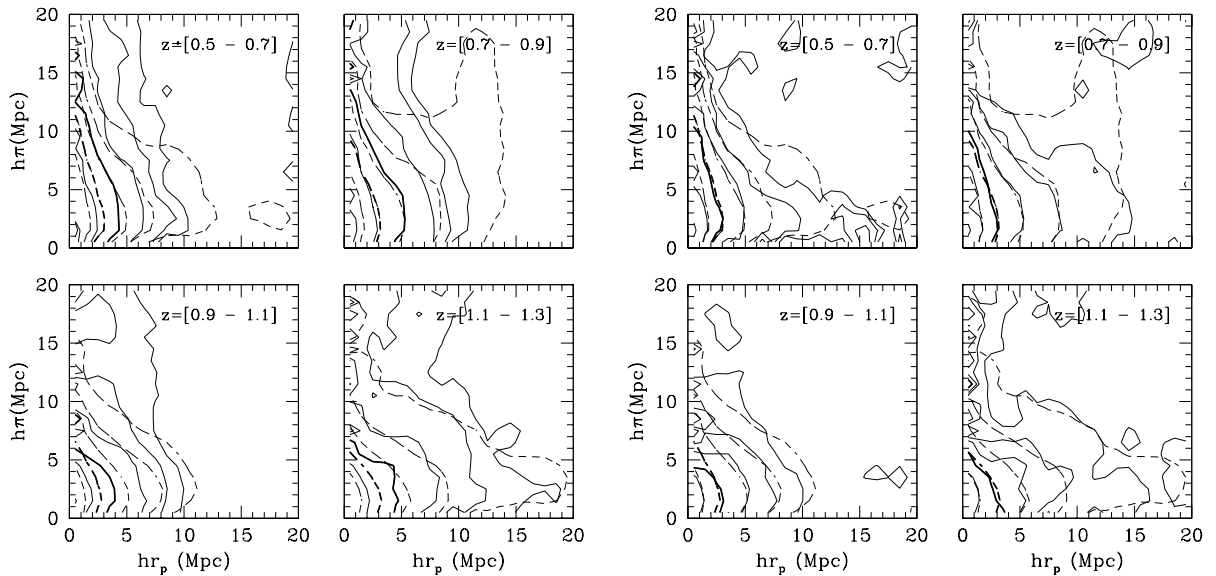


Fig. 10. Same as Fig. 9, but for the $\xi(r_p, \pi)$ correlation function. The contours correspond to values for $\xi(r_p, \pi)$ of 0.4, 1 (bold), 2.0, 5.0. Dashed lines refer to the complete mock sample, while solid ones describe the sample after applying the VVDS selection function.

6.3. Application to redshift-space correlations

We have applied the manipulations presented in the previous section to our mock VVDS 02h surveys and compared the results to those obtained from the whole $1\text{deg} \times 1\text{deg}$ mock fields. The results are shown in Figs. 9–11 for the same mock sample used for measuring $\omega(\theta)$ (Fig. 8), split into 4 redshift bins. In each of these figures, comparison of the the four left to the four right panels demonstrates the effect of the overall correction. In general, in redshift space the effect of the observational biases is much less severe, being diluted by the unaffected clustering measured along the line of sight. Still, we see how a proper estimate does require a correction.

Looking at $\xi(s)$ (Fig. 9), we see that the correction introduced by our scheme is in general very good. The full bi-dimensional correlation function $\xi(r_p, \pi)$ (Fig. 10) shows the effect in more detail, indicating also that the impact of the angular bias on spatial correlations depends on redshift. This is to be expected, given that a fake inhomogeneity at a given angular scale affects larger spatial scales at larger redshifts. However, as seen from the four right panels the bulk of the problem is corrected by our technique.

Finally, the corresponding projected function, $w_p(r_p)$, which is the one that will be fitted to estimate the real-space correlation length and slope, (Le Fèvre et al. 2005), does not show any significant systematic effect, nor scale-dependent

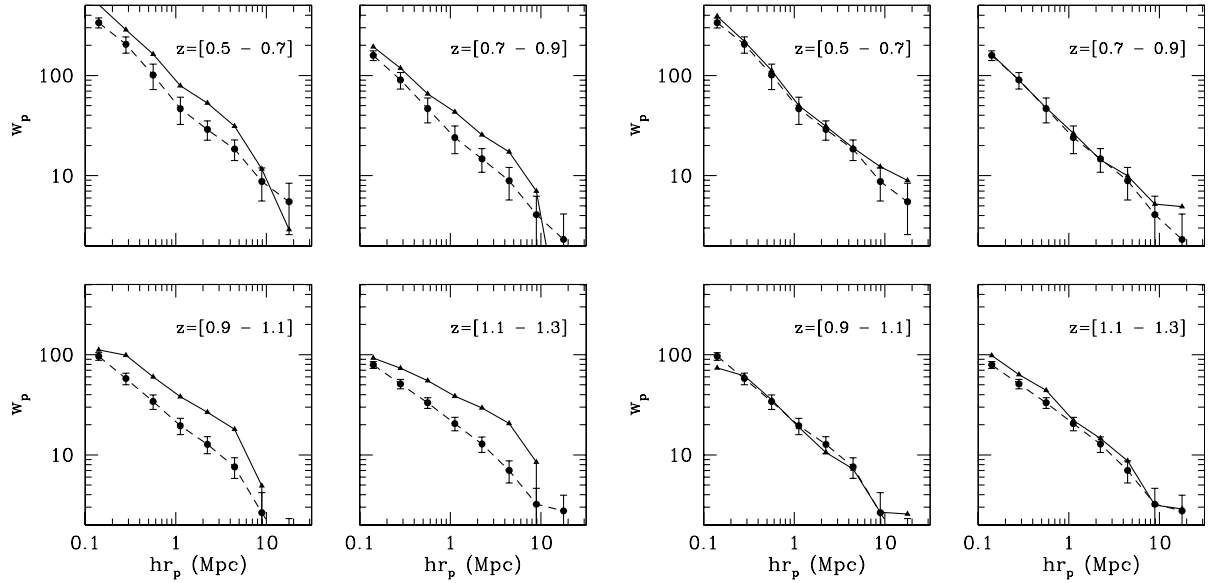


Fig. 11. Same as Figs. 9 and 10, but for the projected function $w_p(r_p)$, measured before (dashed line) and after (solid line) the full observing strategy has been applied. This comparison shows that our method is able to properly recover $w_p(r_p)$. We note, however, that, being closely related to the angular function, $w_p(r_p)$ remains the most sensitive among the 3D correlation functions to the observational biases and the most difficult to recover properly in all r_p bins.

bias (see also Sect. 6.4, below), if one excludes a residual effect in the highest-redshift bin (which may be specific of the mock sample used).

6.4. Accuracy in recovering r_0 and γ

Let us now evaluate more quantitatively how well the weighting scheme is able to recover the correct values of the two parameters of $\xi(r)$, r_0 and γ . Figure 12 plots the projected correlation function w_p , computed for one of the VVDS mock cones, together with the measured best fit values of r_0 and γ . The error contours are estimated from the variance of the 50 mock surveys as described previously and their size depends mainly on the number of galaxies within each bin. Figure 13 shows that the evolution of clustering we “observe” in this specific simulated VVDS cone agrees quite well with its parent sample.

Of course, due to cosmic variance, the values of r_0 and γ differ between different simulated cones. Figure 14 shows the spread of these parameters among all the 50 mock VVDS surveys and their parent catalogs, for a representative redshift bin ($z = [0.5-0.7]$). This behavior is similarly seen in the other redshift bins, indicating an increased spread in the parameter estimates in the “observed” catalogs, an effect easily explained in terms of the smaller number of objects. Figures 12 and 14 also indicate that at the end of our correction process any possible systematic effect is reduced to less than 5%, a value always significantly smaller than the uncertainty due to cosmic variance which is of the order of 15–20%.

6.5. Tests of VVDS observing strategy

In this section we want to discuss from a more general perspective (i.e. not limited to the current status and lay-out of the 02h field) how the accuracy of correlation measurements

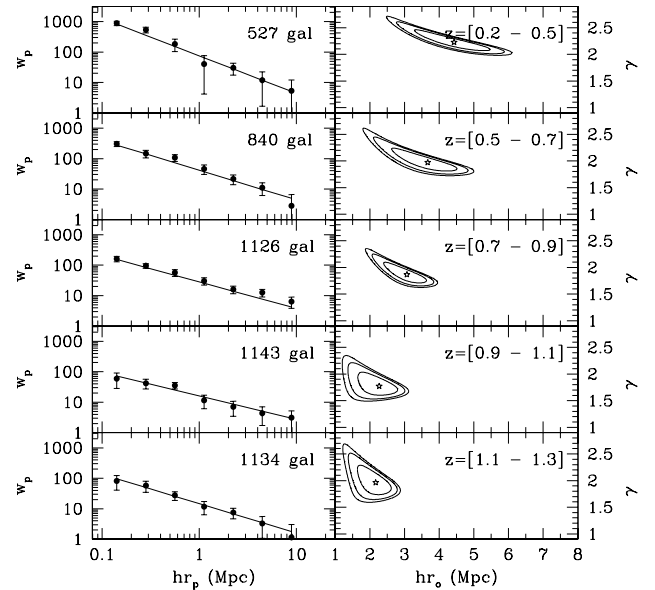


Fig. 12. Evolution of the projected function $w_p(r_p)$ (left column) and the corresponding best-fit parameters of $\xi(r)$, r_0 and γ (right column), as seen in one of the VVDS mock surveys. Error bars are computed as explained in the text, while error contours on the fit parameters are obtained taking into account the full covariance matrix. The 68.3%, 90% and 95.4% joint confidence levels are defined as in Numerical Recipes (Numerical Recipes, Press et al. 1992, chapter 15.6) in terms of the corresponding likelihood intervals that we obtain from our fitting procedure (see Sect. 5.2).

can depend on the number of multiple spectroscopic pointings (“passes”) that are dedicated to a specific area. In other words: are multiple passes increasing – as expected – the accuracy of correlation function measurements, not only thanks to the increased statistics, but also because of the improved sampling

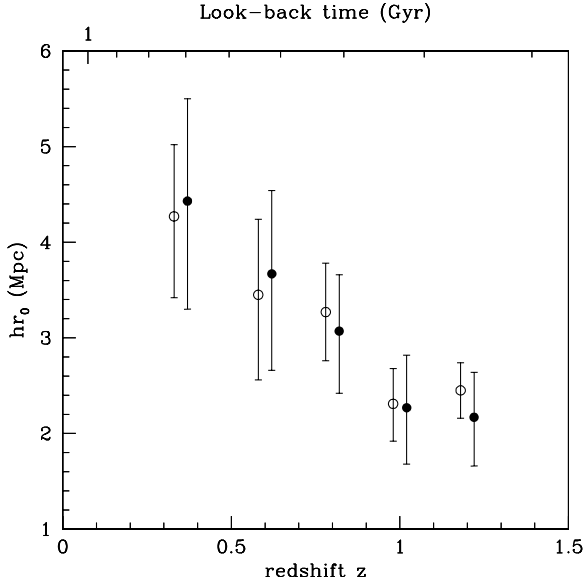


Fig. 13. Evolution of r_0 in a VVDS mock survey (filled circles), compared to that of its parent catalog (open circles). Error bars are as explained in the text. The “true” and “measured” values of r_0 are very consistent within the error bars, providing an *internal* proof of the quality of our correction scheme.

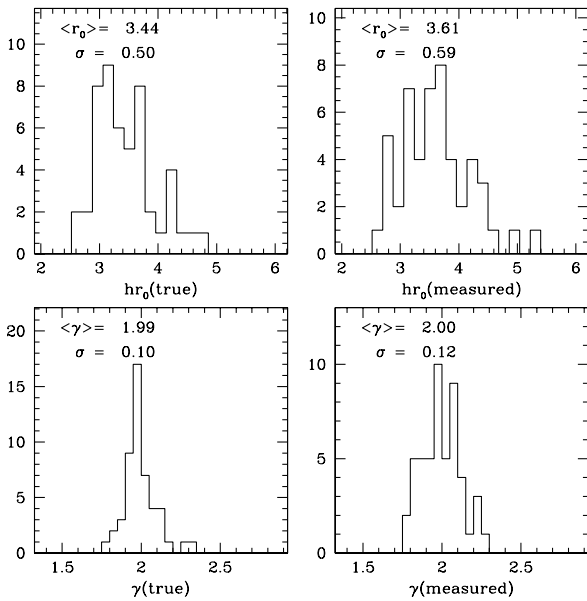


Fig. 14. Histograms of the measurements of r_0 and γ in the redshift bin $[0.5-0.7]$ (chosen as a representative case), among the 50 mock catalogs, for the full cones (*left column*) and for the observed samples (*right column*), where the full weighting scheme has been applied. The ensemble averaged values of r_0 and γ are indicated in each panel, together with their rms error.

of the clustering process? And how is our correcting scheme performing when handling a very sparse (one pass) or a more densely sampled area? This is clearly an interesting question for the future development of the VVDS, or other surveys, as these tests can indicate what strategy could be more efficient. One would like to estimate the fraction of galaxies necessary to recover the correlation signal to a certain level of accuracy.

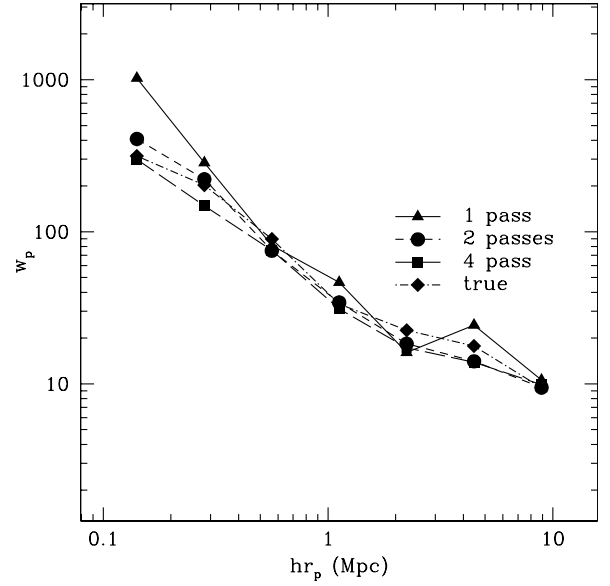


Fig. 15. Measured $w_p(r_p)$ in the case of different number of passes over the same field. When the field is observed only once we are clearly not able to properly recover properly $w_p(r_p)$ on the smallest scale. When we observe the field more times the recovery is much better also on the small scales.

This, translated to the VVDS, implies determining how many spectroscopic “passes” with VIMOS are necessary. Note that the answer is not trivial, since multiple pointings over the same area are usually dithered (i.e. shifted by an amount at least larger than the central “cross”, i.e. $2'$), and thus a larger number of passes over the same area, while improving the sampling, introduces also a more complex mean density pattern, as explained in Sect. 2.1.

Tests have been performed creating a grid of six pointings, spaced with the same step as the real VVDS ones in the VVDS-10h field. The second pass was then arranged over a grid shifted by $2'$ in right ascension and declination. The pointings of both passes have then been “observed” once again with a different selection of objects for spectroscopy. At the end (maximum coverage), this resulted in an area of 0.3624 square degrees, mostly uniformly covered but with small patches of sky that were observed either three, two or one times or remained unobserved. The results for $w_p(r_p)$ and $\xi(s)$ are shown in Figs. 15 and 16, respectively.

The projected correlation function w_p is fairly well recovered almost independently of the sampling density. For a single pass, power is not recovered properly at scales below $0.6 h^{-1}$ Mpc, since there is in practice no pair (even biased) to be “corrected” in a proper statistical way by our scheme.

The case of $\xi(s)$ (Fig. 16) shows even more clearly the difficulty of recovering very small scale pairs with only one pass: in this case, there is an intrinsic low-scale limitation (complete lack of pairs), which cannot be fully overcome by the correcting scheme. The figure shows, for example, that while a linear bin between 0 and $1 h^{-1}$ Mpc is already sufficient to recover the correct clustering amplitude even with one pass, smaller logarithmic bins below $1 h^{-1}$ Mpc are inadequate and suffer from the lack of measured pairs.

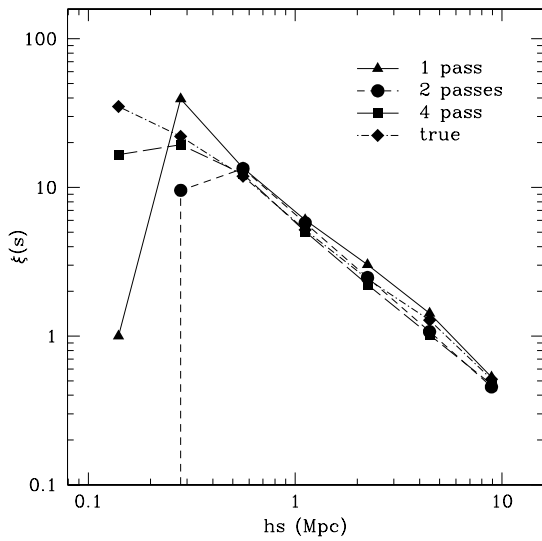


Fig. 16. Measurements of $\xi(s)$ for a different number of observing “visits” over the same field.

We conclude that even in the fields that were observed only with one spectroscopic observation, sampling about 15% of the photometric targets down to $I_{AB} = 24$, the two-point correlation function can be measured quite well for separations $1 \leq r \leq 10 h^{-1}$ Mpc. The results confirm, however, that observing fields four times, sampling about 40% of the population as in the deep part of the VVDS, provides the possibility of more precise measurements on scales down to $0.1 h^{-1}$ Mpc.

7. Summary and conclusions

One of the key goals of the VVDS survey is to measure the evolution of the galaxy clustering from the present epoch up to $z \sim 2$ and larger. To study in detail the error budget of $\xi(r)$ measurements in the VVDS survey, we have generated a set of mock catalogs using the GalICS model of semi-analytic galaxy formation. The geometry of the VVDS survey on the sky is complex due to the observing strategy. The resulting selection function substantially affects the angular correlation properties of the clustering of the observed galaxies. We demonstrate that the correlation observed in redshift space is much less affected and that the bias introduced by the observing strategy can be largely removed using the correcting scheme we propose in this paper.

We conclude that, for the first epoch VVDS data, we can expect to measure $\xi(s)$ and $w_p(r_p)$ to better than 10% on scales $1 \leq r \leq 10 h^{-1}$ Mpc, and better than 30% below $1 h^{-1}$ Mpc. Results obtained from the GalICS simulations indicate that the two-point correlation functions computed from the First Epoch VVDS should suffer only from a modest cosmic variance of $\approx 15\text{--}20\%$. These results suggest that after the final selection of objects for spectroscopy the variance becomes twice as large as the variance of the underlying parent galaxy field in the same area. We expect, in each redshift slice $\Delta z \approx 0.2$ in the redshift range $z = [0.2, 2.1]$, to measure r_0 and γ with an accuracy better than 15–20%. We show that any residual systematic effect in

the measurements of r_0 and γ is below 5%, i.e. a value much smaller than the cosmic errors.

The actual measured clustering properties of galaxies in the VVDS survey, using the framework outlined in this paper, are presented in Le Fèvre et al. (2005b) and in forthcoming papers.

Acknowledgements. We thank the *GalICS* group for access to their simulations, S. Colombi for providing a first set of mock n -body samples early in the development of this work and for useful discussions. This research has been developed within the framework of the VVDS consortium and has been partially supported by the CNRS-INSU and its Programme National de Cosmologie (France), and by Italian Research Ministry (MIUR) grants COFIN2000 (MM02037133) and COFIN2003 (No. 2003020150).

References

- Barrow, J. D., Bhavsar, S. P., & Sonoda, D. H. 1984, *MNRAS*, 210, 19P
- Bernardeau, F., Colombi, S., Gaztañaga, E., & Scoccimarro, R. 2002, *Phys. Rep.*, 367, 1
- Bertin, E., & Arnouts, S. 1996, *A&AS*, 117, 393
- Blaizot, J., Guiderdoni, B., Devriendt, J. E. G., et al. 2004, *MNRAS*, 352, 571
- Blaizot, J., Wadadekar, Y., Guiderdoni B., et al. 2005, *MNRAS*, 360, 159
- Bottini, D., Garilli, B., Maccagni, D., et al. 2005, *A&A*, submitted [arXiv:astro-ph/0409252]
- Eke, V. R., Cole, S., & Frenk, C. S. 1996, *MNRAS*, 282, 263
- Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., & Huchra, J.P. 1994, *MNRAS*, 267, 927
- Giacconi, R., Zirm, A., Wang, J., et al. 2002, *ApJS*, 139, 369
- Guzzo, L., Strauss, M. A., Fisher, K. B., Giovanelli, R., & Haynes, M. P. 1997, *ApJ*, 489, 37
- Hamilton, A. J. S. 1993, *ApJ*, 417, 19
- Hatton, S., Devriendt, J. E. G., Ninin, S., et al. 2003, *MNRAS*, 343, 75
- Hawkins, E., Maddox, S., Cole, S., et al. 2003, *MNRAS*, 346, 78
- Hernquist, L. 1990, *ApJ*, 356, 359
- Kaiser, N. 1986, *MNRAS*, 219, 785
- Kaiser, N. 1987, *MNRAS*, 227, 1
- Landy, S. D., & Szalay, A. S. 1993, *ApJ*, 412, 64
- Le Fèvre, O., Vettolani, G., Maccagni, D., et al. 2003, *The Messenger*, 111, 18
- Le Fèvre, O., Mellier, Y., McCracken, H. J., et al. 2004, *A&A*, 417, 839
- Le Fèvre, O., Vettolani, G., Garilli, B., et al. 2005a, *A&A*, 439, 845
- Le Fèvre, O., Guzzo, L., Meneux, B., et al. 2005b, *A&A*, 439, 877
- Maddox, S. J., Efstathiou, G., Sutherland, W. J., & Loveday, J. 1990, *MNRAS*, 242, 43P
- McCracken, H. J., Radovich, M., Bertin, E., et al. 2003, *A&A*, 410, 17
- Peebles, P. J. E. 1973, *ApJ*, 185, 413
- Peebles, P. J. E. 1980, *The Large Scale Structure of the Universe* (Princeton: Princeton University Press)
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical Recipes in fortran 77: The Art of Scientific Computing*, Second Edition (Cambridge: Cambridge University Press)
- Robin, A. C., Reylé, C., Derrière, S., & Picaud, S. 2003, *A&A*, 409, 523

Chapitre 5

Agrégation des galaxies dans le sondage VVDS

Résumé

5.1 Évolution de la fonction de corrélation globale $\xi(r)$ à $I_{AB} = 24$ jusqu'à $z \sim 2$	127
5.2 Évolution de la fonction de corrélation par type de galaxies jusqu'à $z \sim 2$	140
5.3 Évolution de la fonction de corrélation pour un échantillon de galaxies sélectionnées dans l'ultraviolet	162
5.4 Évolution du biais non-linéaire jusqu'au <i>redshift</i> $z=1.5$	168

Ce chapitre présente les résultats scientifiques sur l'agrégation des galaxies obtenus dans le cadre du sondage VVDS. Je présente tout d'abord les résultats sur l'évolution de la fonction de corrélation globale jusqu'à $z=2.1$, puis les résultats obtenus en divisant l'échantillon par type de galaxies. Je montre ensuite les résultats obtenus sur un sous échantillon de galaxies sélectionnées en ultraviolet.

5.1 Évolution de la fonction de corrélation globale $\xi(r)$ à $I_{AB} = 24$ jusqu'à $z \sim 2$

5.1.1 Données utilisées

Afin d'étudier l'évolution de la fonction de corrélation spatiale $\xi(r)$ jusqu'à redshift $z=2.1$, nous avons utilisé les données des champs *Deep* VVDS-02h et VVDS-CDFS. Les données spectroscopiques utilisées ont été sélectionnées avec le seul critère en magnitude apparente I_{AB} comprise entre 17.5 et 24. Seules ont été conservées pour l'analyse principale les galaxies pour lesquelles la note de qualité du *redshift* mesuré est comprise entre 2 et 9 (niveau de confiance d'au moins $\simeq 75\%$). Les galaxies de note 1 ont été incluses dans un second temps. Les quasars, étoiles (identifiées comme telles) et seconds objets ont également été exclus de l'échantillon final.

L'échantillon final inclus ainsi 6137 galaxies dans le champ VVDS-02h entre $z=0.2$ et 2.1 réparties sur une surface de 1750 minutes d'arc carrées ainsi que 1038 galaxies dans le champ VVDS-CDFS couvrant 453 minutes d'arc carrées entre $z=0.4$ et $z=1.5$.

L'analyse des données du VVDS-02h a été faite dans 6 intervalles en *redshifts* consécutifs ($[0.2 - 0.5]$, $[0.5 - 0.7]$, $[0.7 - 0.9]$, $[0.9 - 1.1]$, $[1.1 - 1.3]$ et $[1.3 - 2.1]$). Celle des données du VVDS-CDFS a été faite uniquement dans 3 intervalles par manque de données ($[0.4 - 0.7]$, $[0.6 - 0.8]$ et $[0.8 - 1.5]$).

Nous avons également appliqué la fonction de sélection en couleur ($R_{AB} \leq 24.1$ et $(B-R) \leq 2.35(R-I) - 0.45$, $(R-I) \geq 1.15$ ou $(B-R) \leq 0.5$) pour une comparaison directe au sondage DEEP2 (Coil *et al.*, 2004) et mesuré les paramètres de la fonction de corrélation dans les mêmes intervalles en *redshifts* que Coil *et al.* (2004) ($[0.7 - 0.9]$, $[0.9 - 1.35]$ et $[0.7 - 1.35]$). Comme le montre la figure 5.1, la fonction de sélection conserve essentiellement des galaxies dont le *redshift* est compris entre ~ 0.7 et ~ 1.35 , conformément à ce qui est attendu, le sondage DEEP2 forçant une présélection pour ne retenir que les galaxies autour de $z \sim 1$: 78% des galaxies de note de confiance 2 à 9 ainsi sélectionnées sont distribuées dans cette plage en *redshift*. Mais une large fraction des galaxies de cette même plage en *redshifts* sont rejetées par cette méthode (51% des notes 1 à 9, 48% des notes 2 à 9). Cela montre l'importance de la fonction de sélection. Le sondage DEEP2, en contraignant à priori la couleur des galaxies pour travailler dans un intervalle de *redshift* restreint, élimine de fait une partie de la population de galaxies dans ce même intervalle de *redshift*. La fonction de corrélation mesurée dans le sondage DEEP2 ne correspond donc pas à la fonction de corrélation de la population globale de galaxies à ces *redshifts*, mais à celle d'un sous échantillon plus rouge en $R - I$ et donc d'avantage corrélié.

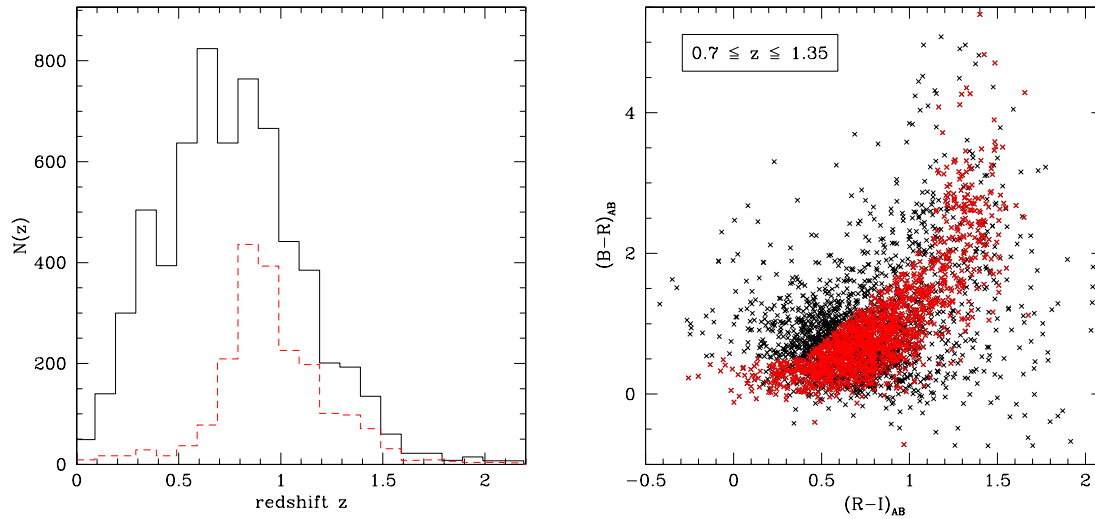


FIG. 5.1 – La figure de gauche reproduit la distribution en *redshifts* des galaxies du VVDS-02h de notes de qualité comprises entre 2 et 9, avec en pointillé rouge la distribution obtenue en appliquant la fonction de sélection du sondage DEEP2. La figure de droite reporte les galaxies situées entre $z=0.7$ et 1.35 avec en rouge celles sélectionnées avec la méthode DEEP2.

5.1.2 Résultats

La figure 5.2 montre la fonction $\xi(r_p, \pi)$ mesurée dans les six intervalles en *redshifts* différents, de $z=0.2$ à $z=2.1$. On a réalisé sur cette figure un lissage sur une échelle de $3 h^{-1} \text{Mpc}$ avec un filtre Gaussien. La distorsion selon la ligne de visée (*Finger-of-God*) est bien visible à tous les *redshifts*, même si le signal est davantage bruité pour les grandes valeurs de z .

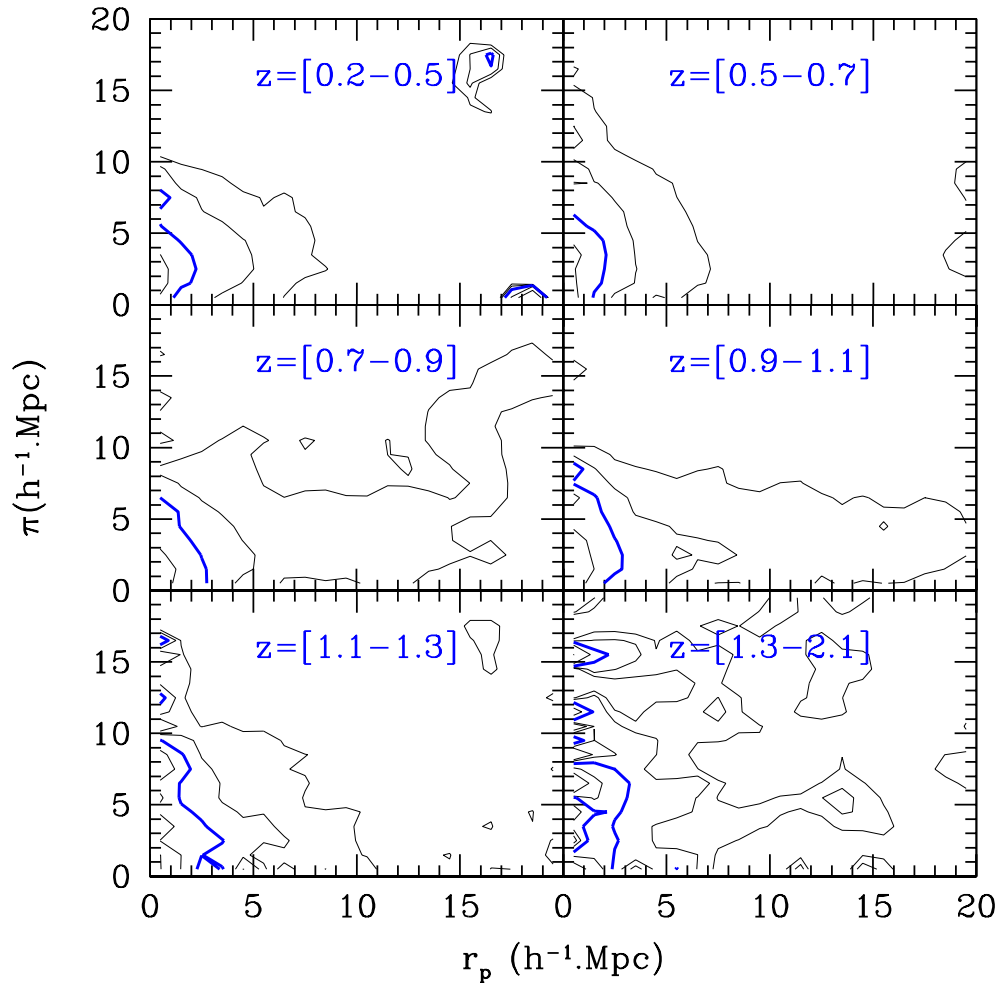


FIG. 5.2 – Fonction $\xi(r_p, \pi)$ pour l'échantillon spectroscopique global du VVDS-02h pour six intervalles en *redshifts* consécutifs. Un lissage sur une échelle de $3 h^{-1} \text{Mpc}$ a été réalisé avec un filtre gaussien. On a représenté ici les niveaux pour $\xi(r_p, \pi) = 0.2, 0.5, 1$ (en gras), 2 et 5

La présentation et l'analyse des résultats de la fonction de corrélation globale sont longuement décrites dans l'article de [Le Fèvre et al. \(2005a\)](#) que l'on trouvera dans cette thèse page 130. On retiendra les points principaux suivant :

- Il s'agit de la première mesure de la fonction de corrélation $\xi(r)$ sur une si grande plage en *redshift* : $[0.2 - 2.1]$
- Un seul critère de sélection : $17.5 \leq l_{AB} \leq 24$

- La longueur de corrélation augmente continuellement entre $r_0=2.23$ et $3.63 \text{ h}^{-1} \text{ Mpc}$ jusqu'à $z \sim 2$
- Les populations sondées sont de plus en plus brillantes avec le *redshift*. L'interprétation de nos mesures n'est donc pas directe.
- Les galaxies observées dans le VVDS à $z \sim 1-2$ sont comparables à des galaxies du 2dFGRS de magnitudes $M_{b_j} = -19.5 + 5 \log(h)$. Nous en déduisons une augmentation de l'amplitude de $\xi(r)$ d'un facteur 2.5 entre $z \sim 1$ et $z=0$.
- Les mesures obtenues en appliquant la stratégie d'observation du sondage DEEP2 sont cohérentes avec celles de [Coil et al. \(2004\)](#) mais les valeurs mesurées ne sont alors pas représentatives de la population principale de galaxies entre $z=0.7$ et 1.35
- Ces observations sont en bon accord avec des résultats de simulations numériques ([Weinberg et al., 2004](#); [Benson et al., 2001](#)). Ils supportent les modèles où les galaxies brillantes tracent les pics principaux du champ de distribution de matière noire à $z \sim 1 - 2$, impliquant une forte évolution du biais.

5.1.3 **Article II : *The VIMOS VLT Deep Survey - The evolution of galaxy clustering to $z \sim 2$ from first epoch observations***

Cet article a été publié dans la revue européenne *Astronomy & Astrophysics* par Le Fèvre, O., Guzzo, L., Meneux, B., Pollo, A. et al. (2005, A&A, 439). Il présente la mesure et l'évolution de la fonction de corrélation spatiale de la population globale des galaxies du sondage VVDS jusqu'à $z \sim 2$.

The VIMOS VLT deep survey

The evolution of galaxy clustering to $z \simeq 2$ from first epoch observations[★]

O. Le Fèvre¹, L. Guzzo², B. Meneux¹, A. Pollo², A. Cappi³, S. Colombi⁹, A. Iovino², C. Marinoni^{1,2},
 H. J. McCracken^{9,12}, R. Scaramella⁷, D. Bottini⁴, B. Garilli⁴, V. Le Brun¹, D. Maccagni⁴, J. P. Picat⁵,
 M. Scodeggio⁴, L. Tresse¹, G. Vettolani⁶, A. Zanichelli⁶, C. Adami¹, M. Arnaboldi¹¹, S. Arnouts¹, S. Bardelli³,
 J. Blaizot¹, M. Bolzonella⁸, S. Charlot^{9,10}, P. Ciliegi⁶, T. Contini⁵, S. Foucaud⁴, P. Franzetti⁴, I. Gavignaud^{5,13},
 O. Ilbert^{1,3}, B. Marano⁸, G. Mathez⁵, A. Mazure¹, R. Merighi³, S. Paltani¹, R. Pellò⁵, L. Pozzetti³, M. Radovich¹¹,
 G. Zamorani³, E. Zucca³, M. Bondi⁶, A. Bongiorno⁸, G. Busarello¹¹, F. Lamareille⁵, Y. Mellier^{9,12},
 P. Merluzzi¹¹, V. Ripepi¹¹, and D. Rizzo^{5,2}

¹ Laboratoire d'Astrophysique de Marseille, UMR 6110 CNRS-Université de Provence, Traverse du Siphon, BP 8,
 13012 Marseille, France
 e-mail: olivier.lefevre@oamp.fr

² INAF – Osservatorio Astronomico di Brera, via Brera 28, 20121 Milan, Italy

³ INAF – Osservatorio Astronomico di Bologna, via Ranzani 1, 40127 Bologna, Italy

⁴ INAF – IASF, via Bassini, 15, 20133 Milano, Italy

⁵ Laboratoire d'Astrophysique de l'Observatoire Midi-Pyrénées, UMR 5572, 14 Av. Ed. Belin, 31400 Toulouse, France

⁶ INAF – Istituto di Radio-Astronomia, via Gobetti 101, 40129 Bologna, Italy

⁷ INAF – Osservatorio Astronomico di Roma, Italy

⁸ Dipartimento di Astronomia, Università di Bologna, via Ranzani 1, 40127 Bologna, Italy

⁹ Institut d'Astrophysique de Paris, UMR 7095, 98bis Bd. Arago, 75014 Paris, France

¹⁰ Max-Planck-Institut für Astrophysik, 85741 Garching, Germany

¹¹ INAF – Osservatorio Astronomico di Capodimonte, via Moiariello 16, 80131 Napoli, Italy

¹² Observatoire de Paris, LERMA, UMR 8112, 61 Av. de l'Observatoire, 75014 Paris, France

¹³ European Southern Observatory, Karl Schwarzschild Str. 2, 85748 Garching bei München, Germany

Received 6 September 2004 / Accepted 11 April 2005

Abstract. This paper presents the evolution of the clustering of the main population of galaxies from $z \simeq 2$ to $z = 0.2$, from the first epoch VIMOS VLT Deep Survey (VVDS), a magnitude limited sample with $17.5 \leq I_{AB} \leq 24$. The sample allows a direct estimate of evolution *from within the same survey* over the time base sampled. We have computed the correlation functions $\xi(r_p, \pi)$ and $w_p(r_p)$, and the correlation length $r_0(z)$, for the VVDS-02h and VVDS-CDFS fields, for a total of 7155 galaxies in a 0.61 deg^2 area. We find that the correlation length in this sample slightly increases from $z = 0.5$ to $z = 1.1$, with $r_0(z) = 2.2\text{--}2.9 h^{-1} \text{ Mpc}$ (comoving), for galaxies comparable in luminosity to the local 2dFGRS and SDSS samples, indicating that the amplitude of the correlation function was $\simeq 2.5$ times lower at $z \simeq 1$ than observed locally. The correlation length in our lowest redshift bin $z = [0.2, 0.5]$ is $r_0 = 2.2 h^{-1} \text{ Mpc}$, lower than for any other population at the same redshift, indicating the low clustering of very low luminosity galaxies, 1.5 mag fainter than in the 2dFGRS or SDSS. The correlation length increases to $r_0 \sim 3.6 h^{-1} \text{ Mpc}$ at higher redshifts $z = [1.3, 2.1]$, as we are observing increasingly brighter galaxies, comparable to galaxies with $M_{B_{AB}} = -20.5$ locally. We compare our measurement to the DEEP2 measurements in the range $z = [0.7, 1.35]$ (Coil et al. 2004, ApJ, in press) and find comparable results when applying the same magnitude and color selection criteria as in their survey. The slowly varying clustering of VVDS galaxies as redshift increases is markedly different from the predicted evolution of the clustering of dark matter, indicating that bright galaxies traced higher density peaks when the large scale structures were emerging from the dark matter distribution 9–10 billion years ago, being supporting evidence for a strong evolution of the galaxy vs. dark matter bias.

Key words. surveys – galaxies: evolution – cosmology: large scale structure of Universe

[★] Based on data obtained with the European Southern Observatory Very Large Telescope, Paranal, Chile, program 070.A-9007(A), and

on data obtained at the Canada-France-Hawaii Telescope, operated by the CNRS of France, CNRC in Canada and the University of Hawaii.

1. Introduction

The evolution of the clustering of galaxies is a key diagnostic element to test the evolution of the universe, allowing direct comparison between observations and theory. In the current paradigm of galaxy formation and evolution, dark matter halos that contain galaxies are expected to merge and grow under the action of gravity. This translates into a continuous evolution of the correlation function $\xi(r, z)$ of dark matter halos, now quite well understood from extensive high resolution numerical simulations (see Weinberg et al. 2004; Benson et al. 2001; Somerville et al. 2001; Kauffmann et al. 1999). As a direct measurement of the space distribution of dark matter halos is not yet feasible, we are compelled to use galaxies as indirect tracers of the dark matter. Unfortunately, as galaxies are complex physical systems, their relationship to the underlying mass, the “bias”, is difficult to estimate. As galaxies and dark matter evolve, the bias may evolve and relating the measurements of galaxy clustering to the evolution of the total mass is not easy, with the bias shown to depend upon galaxy type, luminosity and local environment (Norberg et al. 2002).

The most straightforward indicator of galaxy clustering is the correlation function $\xi(r)$, representing the excess probability over random of finding a galaxy in a given volume, at a fixed distance from another galaxy. The shape and amplitude of the galaxy correlation function at the current epoch is now established to high accuracy. $\xi(r)$ is well described by a power law $\xi(r) = (r/r_0)^{-\gamma}$ over scales $0.1\text{--}10\ h^{-1}\text{ Mpc}$ (Davis & Peebles 1983; Hawkins et al. 2003; Zehavi et al. 2004), with a more refined modelling requiring some extra power over this shape for separations larger than $2\text{--}3\ h^{-1}\text{ Mpc}$ (Guzzo et al. 1991; Zehavi et al. 2004), a feature possibly encoding information on the relation between galaxies and their host dark-matter halos. The local clustering measurements have shown that the correlation length r_0 increases from late types to early type galaxies, from low luminosity to high luminosity and from low to high galaxy density environments (see Giovanelli et al. 1986; Benoist et al. 1996; Guzzo et al. 1997; Norberg et al. 2001), with luminosity being the dominant effect (Norberg et al. 2002). The most recent estimates of the correlation length from the 2dFGRS and SDSS vary from $r_0 = 3$ for late type star-forming galaxies in low density environments to $r_0 = 5\text{--}6$ for galaxies with $M_* = -19.5$, with the clustering amplitude reaching $r_0 = 7.5\ h^{-1}\text{ Mpc}$ for galaxies four times L_* .

At higher redshifts the situation is less clear. Analysis of the projected angular correlation function favors a stable clustering (Postman et al. 1998; Roche et al. 1999; McCracken et al. 2001; Cabanac et al. 2000) but it requires a priori knowledge of the redshift distribution of the galaxy population sampled. A variety of results have been obtained from smaller spectroscopic samples, with comoving correlation lengths r_0 in the range $2\text{--}5\ h^{-1}\text{ Mpc}$ at $z \sim 0.5$ (Le Fèvre et al. 1996; Small et al. 1999; Shepherd et al. 2001; Carlberg et al. 1999) and $r_0 = 3\text{--}5\ h^{-1}\text{ Mpc}$ at $z \sim 3$ (Giavalisco et al. 1998; Foucaud et al. 2003). Recently, first results from the DEEP2 survey have been presented, indicating a correlation length $r_0 = 3.53 \pm 0.81\ h^{-1}\text{ Mpc}$ in $z = [0.7, 0.9]$, and $r_0 = 3.12 \pm 0.72\ h^{-1}\text{ Mpc}$ in $z = [0.9, 1.35]$ (Coil et al. 2004). The main difficulty in

interpreting these results in terms of evolution of the clustering is to relate the population of galaxies observed at a given redshift to a lower redshift population of “descendants”, identified from a well-defined selection function enabling comparisons. Some of these surveys are targeted to specific classes of galaxies, pre-selected via photometric methods, whose relation to the global population is not obvious. The most notable example is represented by galaxies selected via the Lyman-break technique around $z \sim 3$ and $z \sim 4$ (Steidel et al. 1998), which display a clustering strength similar to present-day normal galaxies and therefore represent a very *biased* population, possibly the precursors of giant cluster ellipticals (Governato et al. 1998).

On the other hand, even when selecting purely magnitude-limited samples, one cannot avoid being affected by the complex dependence of clustering on morphology and luminosity evidenced by the wide range of correlation lengths measured in the local Universe, and its evolution as a function of redshift. One needs to observe samples at increasingly high redshifts with the same luminosities, colors (type) and local environments in order to derive the evolution of the clustering of galaxies and hence attempt to derive how the correlation properties of the mass evolve. At high redshifts, the natural observational bias is to sample increasingly brighter and more actively star forming galaxies, which may have a direct impact on our current vision of the evolution.

Finally, high-redshift samples of spectroscopically measured galaxies have been inevitably limited so far to relatively bright objects in small areas on the sky, which contributes to increase the scatter between independent measurements, further complicating their interpretation.

In this paper we present the first attempt to measure the evolution of the clustering in a consistent way across the redshift range $0.2 < z \leq 2.1$, using 7155 galaxies from the VIMOS VLT Deep Survey (VVDS) over more than 0.61 deg^2 . The VVDS is designed to sample the high redshift population of galaxies in the most unbiased way possible, using a simple magnitude selection in the range $17.5 \leq I_{AB} \leq 24$, using several independent fields up to 4 deg^2 each (Le Fèvre et al. 2005). This analysis uses the high-quality First Epoch VVDS sample, which includes 6117 galaxy redshifts in the VVDS-02h (Le Fèvre et al. 2005) and another 1368 galaxies in the VVDS-CDFS (Le Fèvre et al. 2004b) field. We have measured the redshift-space correlation function $\xi(r_p, \pi)$, and computed the projected function $w_p(r_p)$, to recover the value of the galaxy correlation length $r_0(z)$ up to $z \approx 2$, therefore tracing the evolution of the clustering over more than 10 Gy, or 70% of the current age of the universe.

In Sect. 2 we recall the properties of the VVDS First Epoch sample. In Sect. 3, we describe how the correlation function has been computed, referring in large part to the accompanying paper by Pollo et al. (2005) which describes all the methods set up to validate the measurements and compute the errors. In Sect. 4, we present the results in terms of the evolution of $r_0(z)$, and we compare our results to previous measurements where possible. In Sect. 5, we discuss the evolution of the clustering of the global population of galaxies from $z \approx 2$, before concluding in Sect. 6.

This paper is the first in a series of papers to study the clustering of galaxies at high redshift from the VVDS first epoch data. Guzzo et al. (2005, in prep.) will present the clustering evolution from volume limited samples and infer the dependence of clustering upon luminosity; Meneux et al. (2005, in prep.) will present the differences in clustering observed as a function of galaxy type and its evolution and Pollo et al. (2005, in prep.) will investigate the dependence and evolution of clustering as a function of the local environment. Marinoni et al. (2005) and Le Fèvre et al. (2005, in prep.) will look at the evolution of the galaxy – dark matter bias, and subsequent papers will study the clustering from the redshift population $2 \leq z \leq 5$.

We have used a Concordance Cosmology with $\Omega_m = 0.3$, and $\Omega_\Lambda = 0.7$ throughout this paper. The Hubble constant is normally parameterized via $h = H_0/100$, to ease comparison to previous works, while a value $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$ has been used when computing absolute magnitudes. All correlation length values are quoted in comoving coordinates.

2. VVDS first epoch data

2.1. The sample

The VVDS-Deep sample is strictly selected in magnitude in the range $17.5 \leq I_{AB} \leq 24$, from a complete deep photometric survey (Le Fèvre et al. 2004a; McCracken et al. 2003) without any color or shape selection (Le Fèvre et al. 2004b). We have analysed two fields, the VVDS-02h and the VVDS-CDFS. Over the ~ 9600 redshifts measured in these two fields, we have used in the following analysis only objects with a redshift confidence level higher than or equal to $\approx 80\%$ (quality flags 2 to 9 as defined in Le Fèvre et al. 2004b), excluding QSOs. We will only mention briefly below the effect on measured correlations of relaxing the quality threshold, including the poorest redshift measurements (flag 1). The complete sample analysed concerns a total of 7155 galaxies in 2203 arcmin^2 , with 6137 galaxies in the 1750 arcmin^2 VVDS-02h field with $0.2 \leq z \leq 2.1$, and 1038 galaxies in the 453 arcmin^2 VVDS-CDFS (Chandra Deep Field South, Giacomini et al. 2002) area with $0.4 \leq z \leq 1.5$. The accuracy of the redshift measurements is $\approx 275 \text{ km s}^{-1}$ (Le Fèvre et al. 2004a). The distribution of redshifts for the VVDS-02h field is presented in Fig. 1. The redshift distribution peaks at $z \approx 0.8$, and there are 300 galaxies with $1.3 \leq z \leq 1.5$, and 132 galaxies with $1.5 \leq z \leq 2.1$.

2.2. The galaxy population mix

We have split the sample into 6 redshift bins, as described in Table 1. The rest frame $B - I_{AB}(0)$ color and absolute magnitude $M_{B_{AB}}$ distribution within each bin is shown in Fig. 2, from $z = 0.2$ to $z = 2.1$, and the mean values are reported in Table 1. $B - I_{AB}(0)$ and $M_{B_{AB}}$ have been computed using template fitting of the photometric spectral energy distribution in B , V , R , and I bands, to derive the $k(z)$ corrections (see Ilbert et al. 2005, for details).

Up to redshift $z = 1.3$, the $B - I_{AB}(0)$ color distribution stays quite similar with increasing redshift, from blue star forming $B - I_{AB}(0) = 0$ to red $B - I_{AB}(0) \sim 3$, while for $1.3 < z \leq 2.1$,

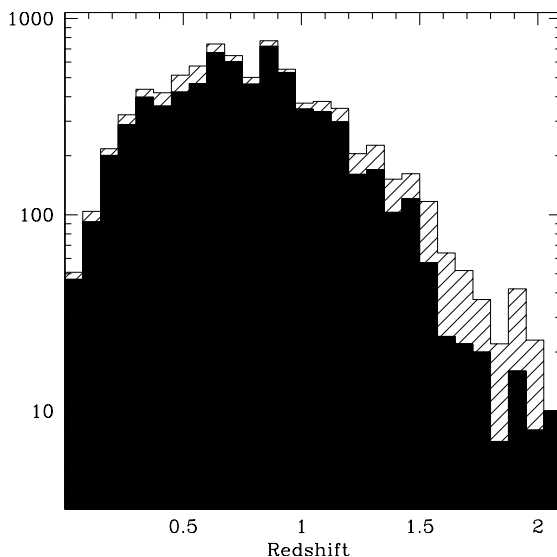


Fig. 1. The redshift distribution in the VVDS-02h field. The filled histogram contains 6137 galaxies with quality flag ≥ 2 , and another 1093 galaxies with quality flag = 1 (open dashed histogram), in the redshift range $z = [0.2, 2.1]$.

the reddest $B - I_{AB}(0) \geq 2.5$ galaxies, if present, are not observed. The magnitude selection of the VVDS-Deep, therefore, allows us to sample the global population of galaxies, for all galaxy types from late to early types, up to $z \sim 1.3$, while for $z > 1.3$, as we are selecting galaxies from their UV rest frame continuum at these redshifts, the VVDS is increasingly biased towards late-type, higher star-forming galaxies. We therefore expect that in the farthest redshift bin, the clustering measurement in the VVDS is the result of the effects of looking at intrinsically more luminous and more actively star forming galaxies.

The range of absolute $M_{B_{AB}}$ magnitudes sampled is quite large, and changes strongly with redshift as shown in Fig. 2. While at redshift $z \sim 0.5$, the absolute magnitude range sampled is $-22 \leq M_{B_{AB}} \leq -16$, only bright galaxies with $M_{B_{AB}} \leq -20$ are sampled at $z > 1.3$. The absolute luminosity of galaxies in the VVDS is shown to increase as redshift increases, as computed from the Luminosity Function (Ilbert et al. 2005). If one assumes a pure luminosity evolution, galaxies observed at high redshift are expected to have faded to fainter absolute magnitudes at the present time, by as much as $M_B = 1.5$ to 2 mag. This should be taken into account when comparing the clustering of the high redshift population to local populations.

The consequences on the correlation function measurements of the change in the population sampled as a function of redshift are discussed in Sect. 5.

3. Computing the real-space correlation and correlation length $r_0(z)$

The methods applied on the VVDS first epoch data to derive the real space correlation parameters are described extensively in the accompanying paper by Pollo et al. (2005). We summarize below the main elements of this method.

Table 1. Measurements of the correlation length $r_0(z)$ and the correlation function slope $\gamma(z)$. The associated 1σ errors are reported. r_0 values are computed both with letting γ free, and setting γ to the mean 1.78 in the range $z = [0.2, 1.3]$ in the VVDS-F02 and VVDS-CDFS. To compare to the DEEP2 measurements of Coil et al. (2004), we have set $\gamma = 1.66$ for the VVDS-F02-DEEP2 in $z = [0.7, 0.9]$ and $z = [0.9, 1.35]$, and $\gamma = 1.68$ for the VVDS-F02-DEEP2 in $z = [0.7, 1.35]$. The redshift range, number of galaxies used, mean absolute magnitude $M_{B_{AB}}$ and rest frame color $B - I_{AB}(z = 0)$, are indicated for each subsample.

Field	Redshift range	N_{gal}	mean $M_{B_{AB}}$	mean $B - I_{AB}(z = 0)$	$r_0(z)$ h^{-1} Mpc	γ	$r_0(z)$ (fixed γ)
VVDS-02h	[0.2–0.5]	1258	−17.57	1.34	$2.23^{+0.46}_{-0.51}$	$1.67^{+0.17}_{-0.11}$	$2.27^{+0.47}_{-0.52}$
VVDS-02h	[0.5–0.7]	1443	−18.68	1.28	$2.69^{+0.53}_{-0.59}$	$1.71^{+0.18}_{-0.11}$	$2.51^{+0.49}_{-0.55}$
VVDS-02h	[0.7–0.9]	1398	−19.20	1.36	$2.87^{+0.51}_{-0.52}$	$1.72^{+0.19}_{-0.12}$	$2.87^{+0.50}_{-0.52}$
VVDS-02h	[0.9–1.1]	1025	−19.66	1.35	$2.87^{+0.62}_{-0.66}$	$1.77^{+0.24}_{-0.14}$	$2.75^{+0.60}_{-0.63}$
VVDS-02h	[1.1–1.3]	561	−20.26	1.35	$3.09^{+0.61}_{-0.65}$	$1.96^{+0.27}_{-0.21}$	$2.93^{+0.58}_{-0.62}$
VVDS-02h	[1.3–2.1]	432	−20.93	1.25	$3.63^{+0.63}_{-0.76}$	$1.92^{+0.31}_{-0.31}$	$3.69^{+0.77}_{-1.00}$
VVDS-02h“DEEP2” selection	[0.7–1.35]	1620	−20.00	1.36	$3.05^{+0.51}_{-0.53}$	$1.56^{+0.14}_{-0.11}$	$3.17^{+0.53}_{-0.55}$
VVDS-02h“DEEP2” selection	[0.7–0.9]	687	−19.65	1.41	$3.59^{+0.77}_{-0.83}$	$1.65^{+0.25}_{-0.17}$	$3.59^{+0.77}_{-0.83}$
VVDS-02h“DEEP2” selection	[0.9–1.35]	933	−20.25	1.32	$3.29^{+0.57}_{-0.59}$	$1.82^{+0.19}_{-0.15}$	$3.13^{+0.54}_{-0.57}$
VVDS-CDFS	[0.4–0.7]	421	−19.44	1.30	$3.19^{+0.93}_{-1.01}$	$1.45^{+0.25}_{-0.17}$	$3.31^{+1.06}_{-1.26}$
VVDS-CDFS	[0.6–0.8]	372	−19.74	1.47	$4.55^{+1.25}_{-1.46}$	$1.48^{+0.28}_{-0.15}$	$4.15^{+1.25}_{-1.50}$
VVDS-CDFS	[0.8–1.5]	452	−20.43	1.34	$3.53^{+1.29}_{-1.66}$	$1.66^{+0.40}_{-0.27}$	$3.49^{+1.28}_{-1.78}$

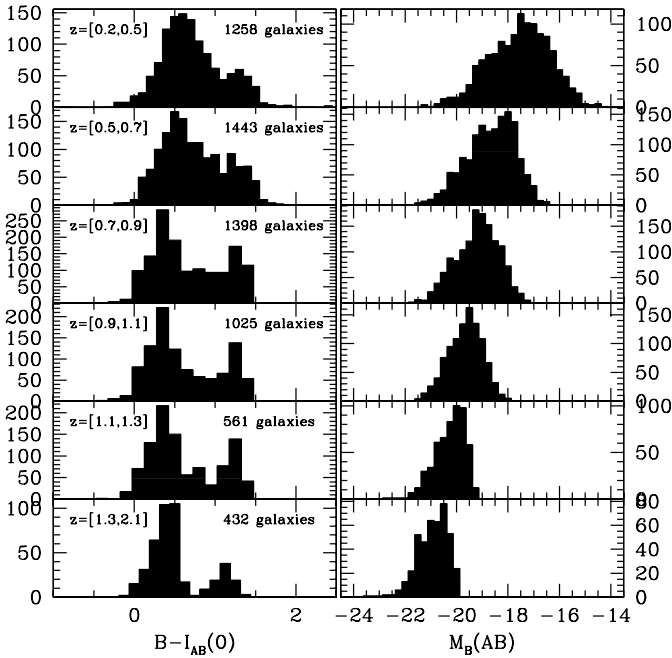


Fig. 2. The rest frame $B - I_{AB}$ color (left) and absolute magnitude $M_{B_{AB}}$ (right) from the VVDS-02h data up to $z \sim 2$.

3.1. Estimating correlation functions from the VVDS

To measure the galaxy real-space correlation length r_0 and slope γ from our survey, we have used the projection of the bi-dimensional correlation function $\xi(r_p, \pi)$. This function was

estimated using the well-known Landy & Szalay (1993) estimator

$$\xi(r_p, \pi) = \frac{N_R(N_R - 1)}{N_G(N_G - 1)} \frac{GG(r_p, \pi)}{RR(r_p, \pi)} - 2 \frac{N_R - 1}{N_G} \frac{GR(r_p, \pi)}{RR(r_p, \pi)} + 1 \quad (1)$$

where N_G is the mean galaxy density (or, equivalently, the total number of objects) in the survey; N_R is the mean density of a catalogue of random points distributed within the same volume of the considered redshift bin; $GG(r)$ is the number of independent galaxy-galaxy pairs with separation between r and $r + dr$; $RR(r)$ is the number of independent random-random pairs within the same interval of separations. A total of $\sim 40\,000$ random points have been used in each redshift bin, guaranteeing a sufficient density to avoid shot-noise effects on small scales. The random sample follows exactly the same geometry, redshift distribution and observational pattern as in the galaxy data, while a specific weighting scheme is used to overcome the biases introduced by the slit-positioning software and other selection effects. These techniques have been extensively tested on a large number of mock VVDS surveys and have been shown to be able to recover the correct $w_p(r_p)$ correlation function to better than 10%, reducing any systematic effect to less than 5%. We have not applied this small correction to the data presented in this paper. This is discussed in detail in the accompanying paper by Pollo et al. (2005).

Since we are not computing the correlation function from the full magnitude-limited survey altogether, there is no point here in using the so-called J_3 minimum-variance weighting. This usually has been adopted in the analysis of large flux-limited local surveys (Fisher et al. 1994; Guzzo et al. 2000),

in which the sampling of the clustering process varies dramatically between the nearby and distant parts of the sample. Its main scope is to avoid excessive weighting of the most distant parts of the sample, where only sparse bright galaxies trace structures. Within each of our redshift bins, the density of objects varies only slightly and equal weighting of the pairs is the most appropriate choice (Fisher et al. 1994).

3.2. $\xi(r_p, \pi)$, $w_p(r_p)$ and the correlation length r_0

We have computed the two point correlation function $\xi(r_p, \pi)$ in increasing redshift bins, selecting the bin boundaries to maximize the number of objects, hence the signal to noise of the correlation function, in each of the bins.

The formalism developed by Davis & Peebles (1983) has been used to derive the real space correlation function in the presence of redshift-space distortions along the line of sight. We integrate $\xi(r_p, \pi)$ along the line of sight, computing

$$w_p(r_p) \equiv 2 \int_0^\infty dy \xi(r_p, \pi) = 2 \int_0^\infty dy \xi \left[(r_p^2 + y^2)^{1/2} \right], \quad (2)$$

where in practice the upper integration limit has to be chosen finite, to avoid adding noise to the result. After a few experiments, we used a value of $20 h^{-1} \text{ Mpc}$.

If we assume a power-law form for $\xi(r)$, i.e.

$$\xi(r) = \left(\frac{r}{r_0} \right)^{-\gamma}, \quad (3)$$

$w_p(r_p)$ can be written as

$$w_p(r_p) = r_p \left(\frac{r_0}{r_p} \right)^\gamma \frac{\Gamma(\frac{1}{2}) \Gamma(\frac{\gamma-1}{2})}{\Gamma(\frac{\gamma}{2})}, \quad (4)$$

where Γ is the Euler Gamma function.

Fitting the $w_p(r_p)$ measurements in each redshift bin then provides a measurement of $r_0(z)$ and $\gamma(z)$.

3.3. Computing errors

The uncertainty associated with the computation of $r_0(z)$ and $\gamma(z)$ is largely dominated by cosmic variance. Although our spectroscopic sample is the largest available to date at the redshifts probed, both in the number of galaxies and area surveyed, only two fields have been sampled, with the VVDS-02h largely dominating over the VVDS-CDFS in terms of the number of galaxies observed and area covered, and it is therefore inappropriate to estimate errors on $w_p(r_p)$, $r_0(z)$, and $\gamma(z)$ directly from field-to-field variations within our data set. Instead, we have been using ensemble errors derived from the scatter in the correlation function computed from 50 mock VVDS surveys constructed using the GalICS simulations (Blaizot et al. 2005), as thoroughly described in the parallel paper by (Pollo et al. 2005). These realistic mock samples specifically include the same selection function and observational biases present in the VVDS First Epoch data set, and allow us to compute ensemble statistical averages and scatters of the quantities we

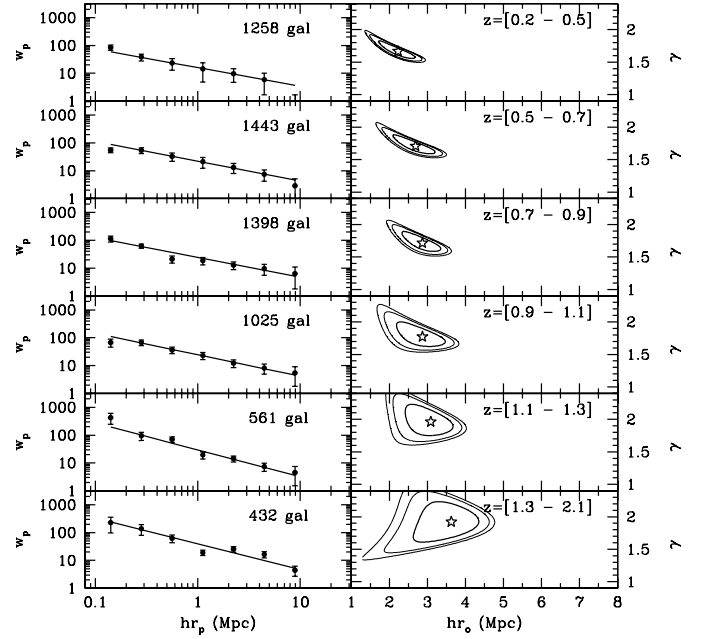


Fig. 3. The correlation function from the VVDS-02h data up to $z \approx 2$: (left) $w_p(r_p)$ measured in 6 redshift bins, the number of galaxies in each bin is indicated on the top right of each panel, (right) the correlation length $r_0(z)$ and correlation function slope γ measured from fitting $w_p(r_p)$ assuming a power law form $\xi(r) = (r/r_0)^{-\gamma}$. 68%, 90% and 95% likelihood contours are drawn.

measure on the real data. In particular, the standard deviation of the measured $w_p(r_p)$ from the mock VVDS samples provides us with a realistic set of error bars that include also the effect of field-to-field variations due to fluctuations on scales larger than the observed field, i.e. cosmic variance (obviously assuming that the Λ -CDM power spectrum and normalization are a good match to the observed one, which is indeed the case).

The comparison of the VVDS-02h and VVDS-CDFS correlation functions gives us an external check to this procedure, with a relatively noisy indication of the amplitude of cosmic variance, as described below.

4. Results

4.1. The VVDS correlation function from a magnitude limited sample $17.5 \leq I_{AB} \leq 24$ out to $z \approx 2$

We have computed the correlation function $\xi(r_p, \pi)$ and its projection $w_p(r_p)$ on both the VVDS-02h and VVDS-CDFS fields.

The correlation function $w_p(r_p)$ is presented in the left panel of Fig. 3 for the VVDS-02h field, and in Fig. 4 for the VVDS-CDFS field. The number of galaxies observed in each field is indicated for each redshift bin in Fig. 3, so far the largest sample of galaxies used to compute the correlation function at these redshifts. The large sample allows to compute $w_p(r_p)$ in 6 redshift bins up to $z = 2.1$ for the VVDS-02h field, and in 3 bins up to $z = 1.5$ for the VVDS-CDFS field. Error bars are ensemble errors computed as described in Sect. 3.3. A positive correlation signal is measured out to at least $30 h^{-1} \text{ Mpc}$ in all bins, and $w_p(r_p)$ is well described by a power law in the range $0.1 \leq r_p \leq 10 h^{-1} \text{ Mpc}$ (note however that any redshift space

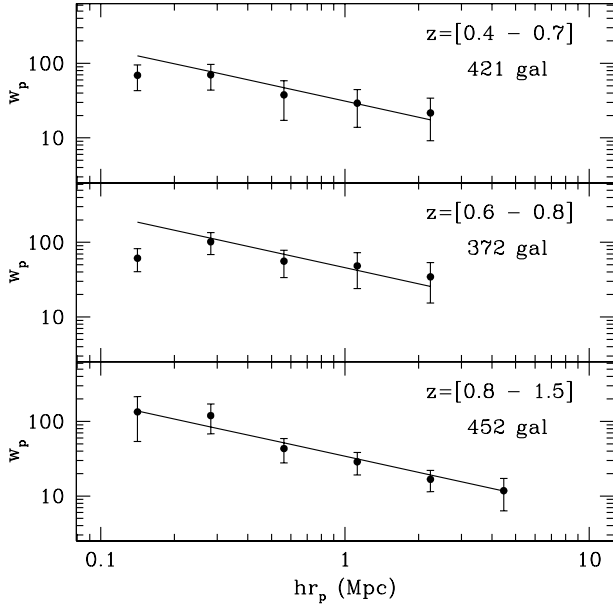


Fig. 4. The correlation function $w_p(r_p)$ for the VVDS-CDFS data in three redshift bins from $z = 0.4$ to $z = 1.5$. The slope has been fixed to $\gamma = 1.72$, i.e. the mean measured in the VVDS-02h field in the range $z = [0.2, 1.1]$.

feature in $\xi(r)$ is smoothed out in $w_p(r_p)$, which is its integral). The measured correlation function amplitude is relatively low at low redshifts $z \leq 0.5$, and stays essentially constant as a function of redshift, with just a mild increase in the farthest bins. A possible interpretation of these results is discussed in the next sections.

4.2. Evolution of the galaxy real-space correlation function

The measured values of the correlation length $r_0(z)$ and the correlation function slope $\gamma(z)$ computed from the fitting of $w_p(r_p)$ are reported in Table 1. For the VVDS-02h field, we have used all w_p points for $0.1 \leq r_p \leq 10 h^{-1}$ Mpc. For the VVDS-CDFS field, we have used points in the range $0.1 \leq r_p \leq 3 h^{-1}$ Mpc only, because of the smaller field size. We also report in Table 1 the values of $r_0(z)$ obtained after fitting w_p with the slope γ fixed to the average slope measured in the range $z = [0.2, 1.3]$. In the VVDS-CDFS, a strong wall-like large scale structure has been identified at $z \sim 0.735$, with more than 130 galaxies in a velocity range $\pm 2000 \text{ km s}^{-1}$ (Le Fèvre et al. 2004b), and is expected to strongly affect the correlation function computation. The correlation function in this field has been computed in the redshift interval $z = [0.6, 0.8]$ which includes this strong structure, as well as in $z = [0.4, 0.7]$ and $z = [0.8, 1.5]$ to specifically exclude it, as reported in Table 1. While we find a slightly higher correlation length in the VVDS-CDFS, the measurements are compatible, within the errors, with the values reported for the VVDS-02h.

We have investigated how our results change when using the full set of available redshifts, i.e. including also the 1300 less accurate measurements (flag 1), which are shown to be $\sim 55\%$ correct (Le Fèvre et al. 2005). The result is that the

measured correlation lengths are lowered by $\sim 3\text{--}5\%$, which we interpret as a consequence of the significant fraction of poorly measured redshifts which dilute the actual correlation function.

The results strikingly show a slowly increasing correlation length over the complete redshift range $z = [0.2, 2.1]$. The lowest value measured is in the lowest redshift bin probed, then r_0 is constant $r_0 \sim 2.8 h^{-1}$ Mpc over the range $z = [0.5, 1.1]$ and increases slightly to $r_0 \sim 3.6 h^{-1}$ Mpc over $z = [1.1, 2.1]$. When we fit the slope γ of the correlation function at the same time as r_0 , it varies between 1.67 and 1.96, slightly increasing with redshift, as reported in Table 1. The average slope in the range $z = [0.2, 1.3]$ is $\gamma = 1.76$.

4.3. Comparison with other surveys

In redshift surveys of the local Universe ($z < 0.2$), the lowest values for the correlation length have been measured for late-type star forming galaxies as e.g. H-II galaxies ($r_0 = 2.7 h^{-1}$ Mpc, Iovino et al. 1988) and infrared-selected *IRAS* galaxies ($r_0 = 3.76 h^{-1}$ Mpc, Fisher et al. 1994). Here we measure an even lower correlation length for galaxies with $z = [0.2, 0.5]$ in the VVDS, coherent with their very low intrinsic luminosity as discussed in Sect. 5.

In order to compare our measurements at higher redshifts to the measurements of the DEEP2 survey (Coil et al. 2004), we have restricted our data by applying a-posteriori the same color and magnitude selection function as the DEEP2 survey applied a-priori to pre-select its spectroscopic targets in the redshift range 0.7–1.35. With this “DEEP2” selection function ($R_{AB} \leq 24.1$, $B - R \leq 2.35(R - I) - 0.45$, $R - I \geq 1.15$ or $B - R \leq 0.5$) only 59% of the VVDS *I*-band magnitude-limited sample galaxies are selected (in the desired redshift range). The correlation function $w_p(r_p)$, and the corresponding $r_0(z)$ and $\gamma(z)$ of this sample are shown in Fig. 5. We find that $r_0 = 3.51 \pm 0.63 h^{-1}$ Mpc for the full $z = [0.7, 1.35]$ sample, compared to 3.19 ± 0.51 by Coil et al. (2004). Separating in the same redshift bins, and setting the slope to $\gamma = 1.66$ as measured in DEEP2, we find that $r_0 = 3.37 \pm 0.67 h^{-1}$ Mpc in $z = [0.7, 0.9]$, $r_0 = 3.05 \pm 0.47 h^{-1}$ Mpc in $z = [0.9, 1.35]$, while Coil et al. (2004) find $r_0 = 3.53 \pm 0.81 h^{-1}$ Mpc and $r_0 = 3.12 \pm 0.72 h^{-1}$ Mpc, respectively, as shown in Fig. 6. Given the relative uncertainties of both surveys, these results are therefore in excellent agreement. The correlation function for the DEEP2 survey therefore differs from the correlation function of the complete VVDS sample because of the different selection function applied that selects galaxies brighter than in the VVDS and excludes galaxies relatively blue in $R - I$ ($R - I < 0.7$) but still red in $B - R$ ($B - R > 0.6$), resulting in a larger correlation length.

5. Discussion: evolution of the clustering length $r_0(z)$ from $z \approx 2$

The evolution of the clustering length $r_0(z)$ from the VVDS first epoch “Deep” sample of galaxies with $17.5 \leq I_{AB} \leq 24$ is presented in Fig. 7. The redshift of each bin is computed as the mean of the redshifts of galaxies in each bin. We find that r_0 is roughly constant or possibly slightly increasing,

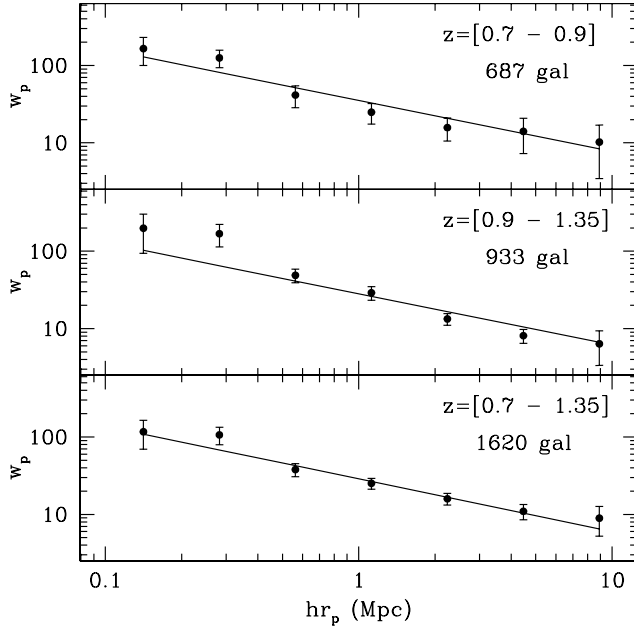


Fig. 5. Correlation function computed from the VVDS sample, applying the same color-magnitude selection function as for the DEEP2 survey (Coil et al. 2004): (*bottom*) the full sample in the redshift range $[0.7, 1.35]$, (*top two panels*) the lower $[0.7, 0.9]$ and higher $[0.9, 1.35]$ redshift samples. The slope has been fixed to the same slope as measured in the DEEP2 to ease r_0 comparison.

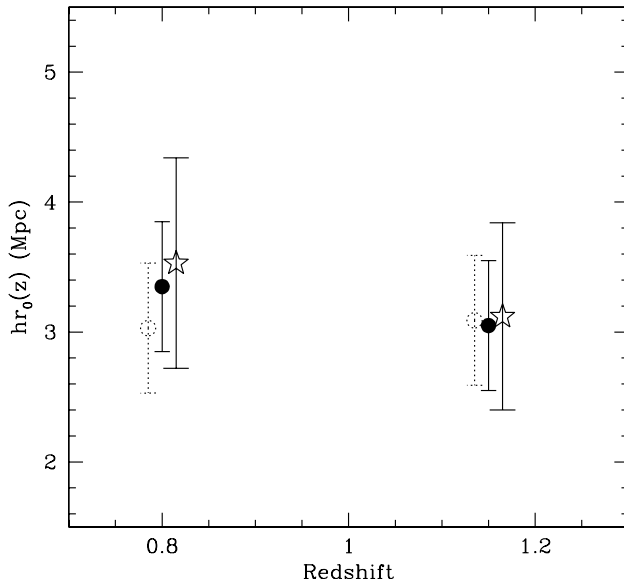


Fig. 6. Comparison of the correlation length $r_0(z)$ (comoving) measured in the VVDS-Deep survey applying a color-color and magnitude selection as in the DEEP2 survey, and the DEEP2 survey measurements (Coil et al. 2004): VVDS-02h, with $\gamma = 1.66$ fixed to the DEEP2 value: filled circles; r_0 and γ fitted simultaneously: open-dot circles, DEEP2 measurements: open stars. Data points have been shifted slightly along the redshift axis to avoid overlap.

within our measurement errors, as redshift increases, with a low value of $r_0 = 2.35^{+0.36}_{-0.37} h^{-1} \text{ Mpc}$ for $z = [0.2, 0.5]$, to $r_0 = 3.03^{+0.51}_{-0.56} h^{-1} \text{ Mpc}$ for $z = [1.3, 2.1]$.

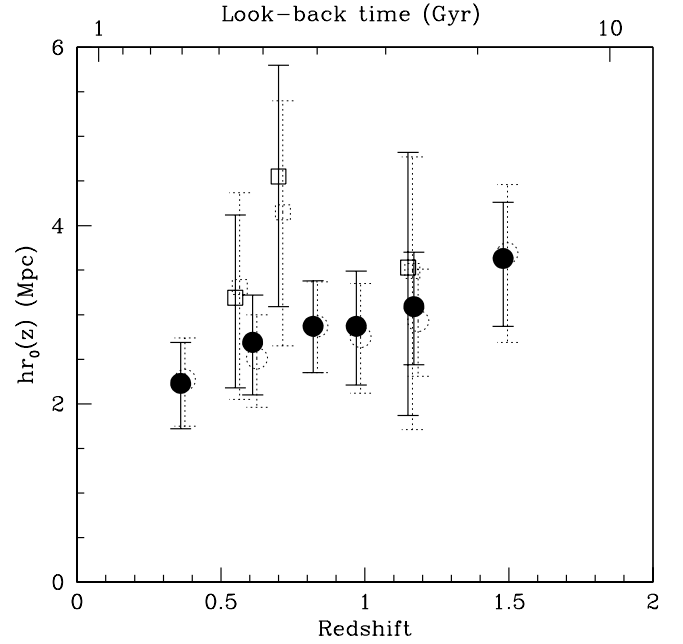


Fig. 7. Evolution of the correlation length $r_0(z)$ (comoving) from the VVDS. Black circles are for the VVDS-02h data, open squares are for the VVDS-CDFS data. The dotted points indicate the measurements when setting the slope to the mean $\gamma = 1.76$ measured in the range $z = [0.2, 1.3]$. Associated errors have been computed from the fitting of $w_p(r_p)$ and associated errors driven from 50 mock galaxy samples from Galics (see text). The VVDS-CDFS measurements are slightly higher than in the VVDS-02h, but remain compatible within the measurement errors. The VVDS-CDFS measurement at $z = 0.7$ include a strong over-density of more than 130 galaxies (see Le Fèvre et al. 2004b), making the correlation function higher in this bin.

Ideally, one would like to follow the evolution of the clustering of the mass in the universe, translated into the gravitational growth of structures. To access to this measurement from the correlation length of galaxies requires understanding of the evolution of galaxies as complex physical systems, and to be able to relate the populations of galaxies observed at different redshifts as descendants of a well-identified original population of matter halos at early epochs. Although the VVDS I -band magnitude selection is the minimum selection bias one can impose selecting a distant sample, this faint magnitude limit leads to a broad redshift coverage, and to a population mix in the VVDS sample that changes with redshift as described in Sect. 2.2. Over the redshift range considered here, the luminosity function of galaxies in the VVDS is strongly evolving up to $\sim 1.5\text{--}2 \text{ mag}$ (Ilbert et al. 2005). The interpretation of the evolution of $r_0(z)$ is therefore not direct.

In the lower redshift bin $z \leq 0.5$, we find a correlation length smaller than any other population of galaxies observed today. The mean absolute luminosity of the low redshift sample is $M_{B_{AB}} = -17.5$ with a significant number of objects fainter than $M_{B_{AB}} = -16$ (Fig. 2), while bright galaxies are under-represented due to the small volume available at these low redshifts. Not surprisingly given the faint $I_{AB} = 24$ cut-off in the VVDS, this makes it the faintest galaxy population for which 3D clustering has ever been probed at low redshifts, about 1.5 mag fainter than the 2dFGRS sample with a mean

$M_{b_j} = -17.98$ equivalent to $M_{B_{AB}} = -18.9$ (Norberg et al. 2001). This low clustering of the low luminosity population of galaxies measured in the VVDS is roughly consistent with an extrapolation to fainter luminosities of the trend for a lower correlation length as galaxies become fainter, as observed locally in the 2dFGRS and SDSS surveys (Norberg et al. 2001; Zehavi et al. 2002), and at intermediate redshifts in the CNOC2 survey (Shepherd et al. 2001).

As redshift increases, the VVDS probes more of the bulk of the general population of galaxies. Galaxies sampled at redshifts 0.5–1.1 become increasingly similar in luminosity and color to the population sampled by the low redshift 2dFGRS and SDSS surveys. As can be seen from Fig. 2, at $z \approx 1$ we measure the clustering of galaxies with a mean absolute magnitude $M_{B_{AB}} = -18.5$, after taking into account ~ 1 mag of luminosity brightening at $z \approx 1$ (Ilbert et al. 2005), thus comparable to the bulk of the galaxies probed by 2dF and SDSS at $z \approx 0.1$.

Over the redshift range $z = [0.5, 1.1]$, we find a slightly increasing correlation length $r_0 \approx 2.2\text{--}2.9 h^{-1} \text{Mpc}$. For a similar, blue-selected, population of galaxies at $z \approx 0.15$, the 2dFGRS finds $r_0 \approx 4.3\text{--}4.6 h^{-1} \text{Mpc}$ (Norberg et al. 2001), using the same technique we use here. Note that at a redshift similar to the 2dFGRS, the SDSS tends to sample a different mix of morphological types, due to its red-based selection, and for this reason measures a larger correlation length, $r_0 = 6.14 \pm 0.18 h^{-1} \text{Mpc}$ (Zehavi et al. 2002; see Hawkins et al. 2003, for discussion). Our results therefore seem to indicate that the amplitude of the correlation function of galaxies which would have luminosities $M_{B_{AB}} = -19.5$ after correction for luminosity evolution is about 2.5 times lower at $z = 1$ than observed locally by the 2dFGRS.

While we know that galaxies with similar luminosities at different redshifts may not be tracing the underlying dark matter in a similar way (see Marinoni et al. 2005), this result is in qualitative agreement with the expectations of the simple gravitational growth of primordial fluctuations (see e.g. Fig. 5 of Weinberg et al. 2004).

In the highest redshift bins $z = [1.1, 1.3]$ and $z = [1.3, 2.1]$, we are measuring the correlation function of the brightest $M_{B_{AB}} \leq -19.5$ galaxies, with a mean of $M_{B_{AB}} \sim -20.5$ and ~ -21.0 respectively. We observe that the correlation length increases compared to the measurements in the range $z = [0.5, 1.1]$, up to $r_0 \approx 3 h^{-1} \text{Mpc}$ at $z = 1.2$ and $r_0 \approx 3.6 h^{-1} \text{Mpc}$ in the highest redshift bin to $z \approx 2$. The luminosity function of galaxies at these redshifts shows a marked evolution, equivalent to an increase in luminosity of ~ 1.5 and ~ 2 mag in $z = [1.1, 1.3]$ and $z = [1.3, 2.1]$, respectively (Ilbert et al. 2005). These galaxies will therefore be expected to have a mean $M_{B_{AB}} \sim -19.5$ at low redshifts after evolution. Again, comparing the clustering observed in the VVDS to the clustering length of $r_0 \approx 5 h^{-1} \text{Mpc}$ measured for $M_{B_{AB}} = -19.5$ in the 2dFGRS (Norberg et al. 2001), we find that the clustering amplitude has increased by a factor ≈ 2.4 from $z \sim 1.3\text{--}1.5$ to $z = 0$.

The evolution of the correlation length observed in our data is in broad agreement with the results of computer simulations of galaxy formation and evolution. In their SPH simulation, Weinberg et al. (2004) find that the clustering length of

galaxies decreases from $r_0 \sim 4.2 h^{-1} \text{Mpc}$ at $z = 3$ to a minimum $r_0 \sim 3.0 h^{-1} \text{Mpc}$ at $z = 1.5$, then increases again to $r_0 \sim 4.0 h^{-1} \text{Mpc}$ at $z = 1$. Note, however, that these predictions refer to the correlation length of the same class of galaxies, corresponding to $M > 5 \times 10^{10} M_\odot$, ideally followed at different redshifts. In a real, magnitude-limited observation, as we have discussed, this effect is related to the different range of luminosities sampled at each redshift, and the changing clustering strength at different luminosities. Besides the simulated clustering of galaxies, the dark matter correlation length is expected from theory and N -body simulations to drop steeply with increasing redshift, from $r_0 \sim 5 h^{-1} \text{Mpc}$ at $z = 0$ to $r_0 \sim 1.8 h^{-1} \text{Mpc}$ at $z = 1.5$ (see e.g. Weinberg et al. 2004). In comparison to these predictions, our data clearly show that the clustering evolution of galaxies does not follow the predicted trend for dark matter.

6. Conclusions

We have computed the evolution of the correlation function $\xi(r_p, \pi)$ and its integral along the line of sight $w_p(r_p)$, from the VVDS first epoch “deep” survey. The VVDS contains a large spectroscopically-selected sample of 7155 galaxies representative of the global galaxy population in the redshift range $z = [0, 2.1]$, in a large 0.61deg^2 total area. The correlation length r_0 is observed to be low, $r_0 \approx 2.2 h^{-1} \text{Mpc}$, for the low redshift $z \leq 0.5$ population, indicating the low clustering of the very low luminosity population sampled in this redshift range. Over the redshift range $z = [0.5, 1.1]$, the correlation length of the population of galaxies, with a luminosity range comparable to the lower redshift 2dFGRS and SDSS, stays roughly constant with $r_0 \approx 2.8 h^{-1} \text{Mpc}$. At the highest redshifts probed in this paper, $z = [1.1, 2.1]$, we find that the correlation length increases slightly to $r_0 \approx 3.6 h^{-1} \text{Mpc}$.

After applying the same selection function as in the DEEP2 survey, our results are found to be in excellent agreement with the results of Coil et al. (2004). However, the significantly different DEEP2 selection function excludes up to 41% of the galaxy population observed by the VVDS, making the correlation length of the DEEP2 sample larger than in the purely magnitude-selected VVDS. The VVDS results demonstrate that the clustering of the whole galaxy population at the same redshift is indeed significantly lower, with $r_0 \approx 2.9$ in the VVDS vs. $r_0 \approx 3.5 h^{-1} \text{Mpc}$ in the DEEP2, further emphasizing the importance of our simple, purely magnitude-limited selection function.

Our measurements clearly show that the correlation length evolves only slowly with redshift in the range $0.5 \leq z \leq 2$, in a magnitude limited sample with $17.5 \leq I_{AB} \leq 24$. Taking into account the different VVDS galaxy populations probed as a function of redshift, with intrinsically brighter galaxies probed as redshift increases, we find that the clustering of galaxies at $z \sim 1\text{--}2$ in the VVDS is about 2.5 times lower in amplitude than for the galaxies probed by the 2dFGRS at $z \sim 0.15$, for populations with similar absolute M_B magnitudes. This result provides unambiguous evidence for clustering evolution.

Our results are in broad agreement with simulations accounting for both gravitational growth and baryonic physics

(Weinberg et al. 2004; Benson et al. 2001). These simulations show that the underlying dark matter correlation evolves strongly with redshift, as expected in a hierarchical growth of structures. Our observation that the clustering of galaxies does not follow such a strong evolution therefore fully supports the model in which luminous galaxies at $z = 1-2$ (or 9–10 billion years ago) trace the emerging peaks of the large-scale dark-matter distribution and implies a strongly evolving galaxy bias.

We will investigate the evolution of the galaxy – dark matter bias elsewhere (Marinoni et al. 2005). A more detailed study of the dependence of clustering on luminosity and galaxy type will be presented in forthcoming papers.

Acknowledgements. This research has been developed within the framework of the VVDS consortium (formerly VIRMOS consortium).

This work has been partially supported by the CNRS-INSU and its Programme National de Cosmologie (France), and by Italian Research Ministry (MIUR) grants COFIN2000 (MM02037133) and COFIN2003 (No. 2003020150).

The VLT-VIMOS observations have been carried out on guaranteed time (GTO) allocated by the European Southern Observatory (ESO) to the VIRMOS consortium, under a contractual agreement between the Centre National de la Recherche Scientifique of France, heading a consortium of French and Italian institutes, and ESO, to design, manufacture and test the VIMOS instrument.

References

- Benoist, C., Maurogordato, S., da Costa, L. N., Cappi, A., & Schaeffer, R. 1996, *ApJ*, 472, 452
- Benson, A. J., Frenk, C. S., Baugh, C. M., Cole, S., & Lacey, C. G. 2001, *MNRAS*, 327, 1041
- Blaizot, J., Wadadekar, Y., Guiderdoni, B., et al. 2005, *MNRAS*, 360, 159
- Cabanac, R. A., de Lapparent, V., & Hickson, P. 2000, *A&A*, 364, 349
- Carlberg, R., Yee, H. K. C., Morris, S. L., et al. 1999, *Phil. Trans. R. Soc. Lond. A*, 357, 167
- Coil, A. L., Davis, M., Madgwick, D. S., et al. 2004, *ApJ*, in press
- Colless, M. M., Dalton, G., Maddox, S., et al. 2001, *MNRAS*, 328, 1039
- Davis, M., & Peebles, P. J. E. 1983, *ApJ*, 267, 465
- Fisher, K. B., Davis, M., Strauss, M. A., Giovanelli, R., & Haynes, M. P. 1994, *MNRAS*, 266, 50
- Foucaud, S., McCracken, H. J., Le Fèvre, O., et al. 2003, *A&A*, 409, 835
- Giacconi, R., Zirm, A., Wang, J., et al. 2002, *ApJS*, 139, 369
- Gialalisco, M., Steidel, C. C., Adelberger, K. L., et al. 1998, *ApJ*, 503, 543
- Giovanelli, R., Haynes, M. P., & Chincarini, G. 1986, *ApJ*, 300, 77
- Governato, F., Baugh, C. M., Frenk, C. S., et al. 1998, *Nature*, 392, 359
- Guzzo, L., Iovino, A., Chincarini, G., Giovanelli, R., & Haynes, M. P. 1991, *ApJ*, 382, L5
- Guzzo, L., Strauss, M. A., Fisher, K. B., Giovanelli, R., & Haynes, M. P. 1997, *ApJ*, 489, 37
- Guzzo, L., Bartlett, J. G., Cappi, A., et al. 2000, *AA*, 355, 1
- Hawkins, E., Maddox, S., Cole, S., et al. (2dF Team) 2003, *MNRAS*, 346, 78
- Ilbert, O., Tresse, L., Zucca, E., and the VVDS consortium 2005, *A&A*, 439, 863
- Iovino, A., Melnick, J., & Shaver, P. 1988, *ApJ*, 330, L17
- Kauffmann, G., Colberg, J. M., Diaferio, A., & White, S. D. M. 1999, *MNRAS*, 303, 188
- Landy, S. D., & Szalay, A. S. 1993, *ApJ*, 412, 64
- Le Fèvre, O., Hudon, D., Lilly, S. J., et al. 1996, *ApJ*, 461, 534
- Le Fèvre, O., and the VVDS consortium 2003, *The Messenger* 111, 18
- Le Fèvre, O., Mellier, Y., McCracken, H. J., et al. 2004a, *A&A*, 417, 839
- Le Fèvre, O., and the VVDS consortium 2004b, *A&A*, 428, 1043
- Le Fèvre, O., Vettolani, G., Garilli, B., and the VVDS consortium 2005, *A&A*, 439, 845
- Lilly, S. J., Le Fèvre, O., Crampton, D., Hammer, F., & Tresse, L. 1995, *ApJ*, 455, 50
- McCracken, H. J., Le Fèvre, O., Brodwin, M., et al. 2001, *A&A*, 376, 756
- McCracken, H. J., Radovich, M., Bertin, E., et al. 2003, *A&A*, 410, 17
- Marinoni, C., Le Fèvre, O., Meneux, B., and the VVDS consortium, *A&A*, submitted
- Norberg, P., Baugh, C. M., Hawkins, E., et al. 2001, *MNRAS*, 328, 64
- Norberg, P., Baugh, C. M., Hawkins, E., et al. 2002, *MNRAS*, 332, 827
- Pollo, A., Meneux, B., Guzzo, L., and the VVDS consortium 2005, *A&A*, 439, 887
- Peacock, J. 2002, Tenerife Winter School, Dark matter and dark energy in the universe [[arXiv:astro-ph/0309240](https://arxiv.org/abs/astro-ph/0309240)]
- Postman, M., Lauer, T. R., Szapudi, I., & Oegerle, W. 1998, *ApJ*, 506, 33
- Roche, N., & Eales, S. A. 1999, *MNRAS*, 307, 703
- Shepherd, C. W., Carlberg, R. G., Yee, H. K. C., et al. 2001, *ApJ*, 560, 72
- Small, T. A., Ma, C., Sargent, W. L. W., & Hamilton, D. 1999, *ApJ*, 524, 31
- Somerville, R. S., Lemson, G., Sigad, Y., et al. 2001, *MNRAS*, 320, 289
- Schneider, D. P., Fan, X., Hall, P. B., et al. 2003, *AJ*, 126, 2579
- Steidel, C. C., Adelberger, K. L., Dickinson, M., et al. 1998, *ApJ*, 492, 428
- Weinberg, D. H., Davé, R., Katz, N., & Hernquist, L. 2004, *ApJ*, 601, 1
- Zehavi, I., Blanton, M. R., Frieman, J. A., et al. 2002, *ApJ*, 571, 172
- Zehavi, I., Weinberg, D. H., Zheng, Z., et al. 2004, *ApJ*, 608, 16

5.2 Évolution de la fonction de corrélation par type de galaxies jusqu'à $z \sim 2$

Nous avons fait état dans les parties 1.5.1.4 et 1.5.1.5 de nombreuses études menées dans l'Univers proche sur l'agrégation des galaxies en fonction de leur type morphologique ou spectral. Elles mettent en évidence une ségrégation claire : les galaxies les plus précoces sont plus agrégées que les types tardifs. Peu d'études à grands *redshifts* ont été menées pour étudier l'évolution de cette différence fondamentale. Les mesures réalisées à $z \sim 1$ et $z \sim 3$ concernent en effet des classes d'objets présélectionnés (*Extremely Red Objects* et *Lyman-Break*). Une classification en type de l'échantillon complet à $I_{AB} = 24$ du VVDS permet pour la première fois de suivre l'agrégation des galaxies en fonction du *redshift* jusqu'à $z \sim 2$ et ainsi de suivre les mécanismes de formation des galaxies. Dans un modèle hiérarchique, les galaxies spirales se seraient formées suite au refroidissement d'un disque de gaz froid alors que les elliptiques seraient le résultat de fusion de plusieurs galaxies (section 1).

Nous nous focalisons ici sur le champ profond VVDS-02h, le seul champ du VVDS qui contient aujourd'hui suffisamment de galaxies pour permettre une subdivision par types de galaxies.

5.2.1 Définition des types

Plusieurs approches existent pour définir une classification d'un échantillon de galaxies :

- classification à l'oeil
- mesure automatique de la forme de l'objet
- décomposition du spectre sur une base de vecteurs propres (*Principal Component Analysis*)
- mesure de la largeur équivalente de certaines raies spectrales
- ajustement des données photométriques multi-couleur à des spectres de référence

La classification utilisée ici dans le cadre du VVDS est basée sur cette dernière technique. Elle consiste à associer chaque galaxie du VVDS à une distribution d'énergie spectrale (SED) de référence en fonction de ces couleurs obtenues à partir des magnitudes apparentes en bandes B, V, R, I, et de son *redshift*. La bande U, qui ne couvre pas tout le champ du VVDS-02h, n'a pas été utilisée afin d'obtenir une classification homogène spatialement. Cette méthode a été utilisée par Zucca *et al.* (2005) dans le cadre de la mesure de la fonction de luminosité par type dans le VVDS et largement développée dans la thèse d'Ilbert (2004).

Les SEDs de références sont les 4 spectres CWW (Coleman *et al.*, 1980, voir figure 1.10) auxquels on rajoute deux modèles de galaxies à fort taux de formation d'étoiles (*starburst*) issus de la librairie GISSEL (Bruzual et Charlot, 1993). Ce jeu de *templates* présente l'avantage d'être simple et standard. De nombreux auteurs l'utilisent en effet afin de classer leur jeu de données (Shepherd *et al.*, 2001; Firth *et al.*, 2002; Budavári *et al.*, 2003, par exemple).

Les propriétés spectro-photométriques de ces *templates* sont utilisées pour classer les galaxies du VVDS quelque soit leurs *redshifts* bien que les modèles montrent que ces

propriétés évoluent au cours du temps. L'utilisation de *templates* synthétiques prenant en compte cette évolution ne permet pas d'obtenir une classification robuste et facile à interpréter à cause de multiples dégénérescences possibles entre la couleur, l'âge, l'extinction, la métallicité et l'histoire de formation d'étoiles. De plus l'interprétation des résultats est plus directe s'ils ne sont pas dépendant d'un modèle. Notre classification ne permet pas de suivre directement l'évolution d'une certaine classe d'objets. Par contre, à une époque donnée, les galaxies d'une même classe présentent statistiquement les mêmes propriétés spectro-photométriques. En ce sens, notre classification peut aussi s'apparenter à une classification en couleur dans laquelle les galaxies les plus tardives sont les plus bleues.

Quatre classes de galaxies sont donc définies. Les types 1 ont des propriétés spectro-photométriques de galaxies elliptiques avec une population dominante d'étoiles vieilles, alors que les types 4 sont associées à des galaxies irrégulières à fort taux de formation stellaire. Le tableau 5.1 résume leurs propriétés dans le cadre du VVDS-02h alors que la figure 5.3 montre leur distribution en *redshift*.

TAB. 5.1 – Propriété des 4 types de galaxies du champs VVDS-02 (notes de confiance des *redshifts* de 2 à 9)

type		nombre de galaxies	couleur indicative au repos
1	E/S0	645	$1.3 < (B - I)_{AB}$
2	S _{bc}	1004	$0.95 < (B - I)_{AB} < 1.3$
3	S _{cd}	2104	$0.68 < (B - I)_{AB} < 0.95$
4	Irrégulière	2742	$(B - I)_{AB} < 0.68$

5.2.2 Mesures et erreurs associées

Comme dans le cas de la fonction de corrélation globale, on calcule la fonction $\xi(r_p, \pi)$ et sa projection $w_p(r_p)$ dans différents intervalles en *redshifts*. Par contre, il est bon de préciser certains points dont il est important de tenir compte.

Calcul du poids associé à chaque galaxie

La méthode pour corriger des effets de sélection et de la stratégie d'observation consiste à attribuer un poids à chaque galaxie de l'échantillon spectroscopique utilisé lors du comptage des paires séparées d'une distance r . Elle est décrite dans la partie 3.3.2. Ce calcul de poids est réalisé en considérant localement le catalogue photométrique de référence afin d'estimer le nombre de paires manquantes qu'il faut corriger. Il est important ici de garder à l'esprit que ce catalogue doit avoir les mêmes propriétés que celles du catalogue spectroscopique. Il correspond en quelque sorte à la convolution entre le catalogue spectroscopique dont on disposerait si le *redshift* de tous les objets étaient connus par la fonction de sélection. Dans le cas ici présent, cette fonction de sélection sera par exemple :

$$\begin{cases} 17.5 \leq I_{AB} \leq 24 \\ \text{galaxies de type 1} \end{cases} \quad (5.1)$$

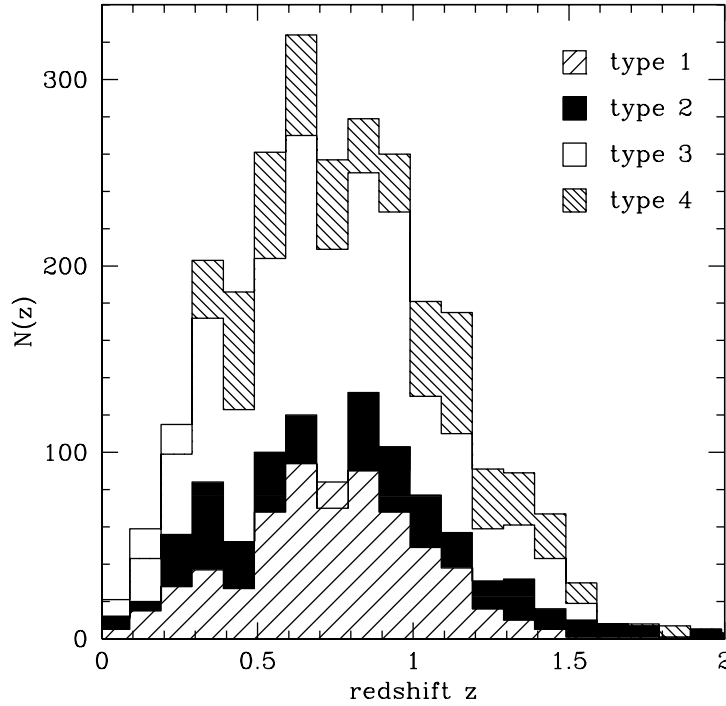


FIG. 5.3 – Distributions en *redshifts* des 4 types de galaxies du VVDS-02h

Dans cet exemple, le catalogue photométrique de référence ne doit donc contenir que des galaxies de type 1. Le poids que l'on cherche à calculer serait modifié par la présence d'une galaxie de type 4, alors que ces deux populations ont des propriétés d'agrégation différentes. Ainsi, nous utilisons les *redshifts* photométriques (calculés à partir des bandes u, g, r, i et z du CFHTLS et B, V, R, I du VVDS) ainsi que les magnitudes apparentes B, V, R, I pour attribuer un type à tous les objets du catalogue photométrique de référence à $17.5 \leq I_{AB} \leq 24$. On obtient ainsi un catalogue de référence pour chaque type.

On notera que les *redshifts* photométriques ne sont utilisés que pour attribuer un type aux objets du catalogue photométrique. Ils ne sont pas directement utilisés dans le calcul de la fonction de corrélation.

Estimation des erreurs

Dans le cas de la fonction de corrélation globale, l'estimation des erreurs a été obtenue à partir de la variance mesurée sur les 50 champs *GalICS*, dans les différents intervalles en *redshifts* considérés. Dans un premier temps j'ai exploré la possibilité de garder la même approche. Pour cela, la base de donnée *GalICS* fournissant les *redshifts* et magnitudes apparentes dans les mêmes filtres B, V, R et I de la caméra CFH12k utilisée pour la photométrie du VVDS, il est possible d'appliquer la même procédure pour attribuer un type que celle définie précédemment.

Malheureusement, la simulation *GalICS* n'est pas aussi réaliste qu'on le voudrait et ne reproduit pas la distribution des types de galaxies observée dans l'Univers lointain. Ainsi,

la figure 5.4 reproduit les distributions en *redshifts* obtenues à partir de notre méthode pour chacun des quatre types. Sur cette même figure se trouvent superposées les distributions de chaque type obtenues dans le VVDS (voir figure 5.3). On constate que notre méthode, qui se base sur les magnitudes observées en BVRI notamment et le *redshift*, n'identifie aucune galaxie précoce de type 1 au-delà de $z \sim 0.5$ dans les simulations. Il sera donc impossible d'estimer la moindre erreur de mesure pour cette population au-delà de cette limite.

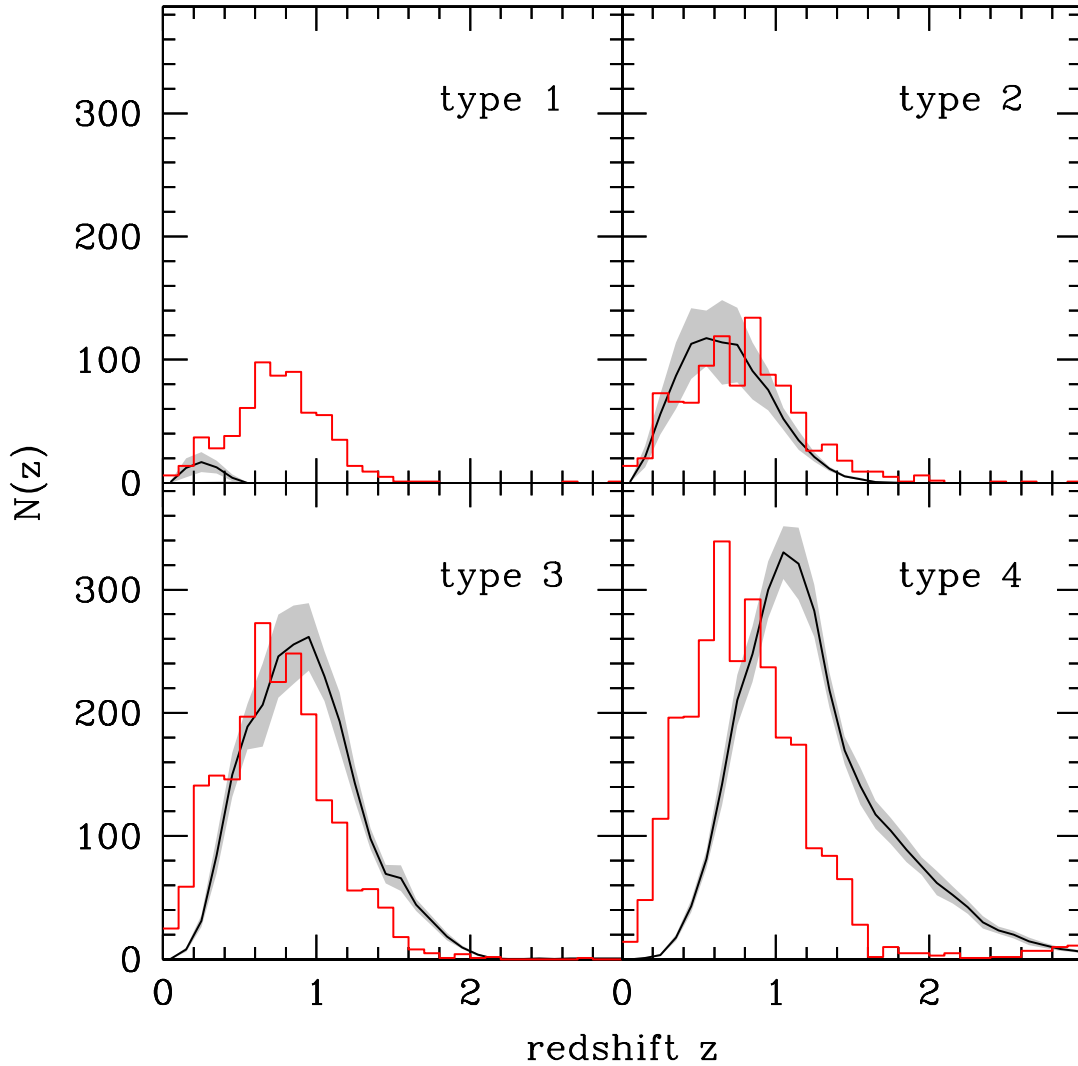


FIG. 5.4 – Distributions en *redshifts* des quatre types de galaxies définis dans 20 champs *GalICS* selon la méthode de Zucca *et al.* (2005). En rouge, on a reporté les distributions en *redshifts* des quatre types identifiés de la même manière dans le VVDS. Aucun objet des simulations *GalICS* n'est identifié comme type 1 au-delà de $z \sim 0.5$.

De plus, les types fournis dans la base de donnée *GalICS* ne permettent pas non plus de définir quatre classes équivalentes aux nôtres (figure 5.5). La base de donnée

GalICS fournit une classification “morphologique” de chaque galaxie. Cette classification est basée sur le rapport des luminosités mesurées en bande B au repos du bulbe et du disque de chaque galaxie. Les types E et S0 regroupent ainsi les galaxies dominées par leur bulbe et les types spirales, Sp, celles dominées par leur disque. Une galaxie est classée comme “Immature” si l’âge de son plus vieux progéniteur est inférieur à 1 million d’années. Et enfin, la classe “Particulière” rassemble les quelques objets pour lesquels ce rapport diverge. La figure 5.5 montre la distribution en *redshift* des objets *GalICS* répartis suivant cette classification. On a regroupé les classe E et S0 d’un coté, Sp, Immature et Particulière de l’autre. Cette classification ne permet pas, là encore, de définir quatre classes équivalentes aux nôtres. Notre type 1 pourrait être associé à la réunion des types E et S0 de *GalICS* comme l’indique la figure 5.5 mais il devient difficile de diviser en 3 la classe de spirales qui représente en moyenne près de 84% de l’ensemble des galaxies *GalICS* (les classes “Immatures” et “Particulière” sont marginales en terme de statistique et ne représente en moyenne à elles deux que 2% de l’ensemble).

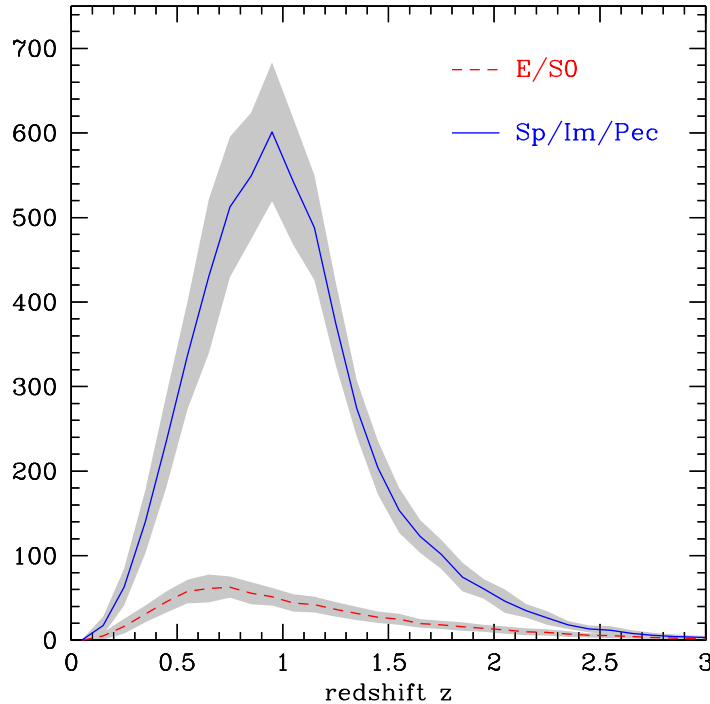


FIG. 5.5 – Moyenne des distributions en *redshifts* des types E/S0 et Spirales/Irrégulières ainsi qu’ils le sont définies dans la base de donnée *GalICS*

Les erreurs obtenues à partir du rééchantillonnage *bootstrap* dans le cas de la fonction de corrélation globale sous-estime d’un facteur $\sim 1.5-2$ les erreurs finales sur r_0 et γ obtenues à partir des simulations *GalICS*. On s’attend donc à avoir un facteur du même ordre de grandeur ici.

L’approche qui a été choisie finalement pour quantifier les erreurs associées aux fonctions de corrélation mesurées pour chaque type consiste à tirer aléatoirement dans chaque champ *GalICS* la même proportion de galaxies d’un type donné que celle observée dans

les données du VVDS, ceci en fonction du *redshift*. Même si la longueur de corrélation r_0 alors mesurée est celle de la population globale, avec des incertitudes de mesures plus grandes, chaque population est correctement représentée d'un point de vue statistique. La variance de champ à champ est artificiellement augmenté du fait que l'on mesure une large plage de valeurs pour r_0 . Ainsi cette méthode se veut très conservative, quitte à surestimer les erreurs.

5.2.3 Résultats

J'ai calculé les fonctions de corrélation $\xi(r_p, \pi)$ pour chacun des 4 types ainsi que pour la réunion des types 1+2 et 3+4. La figure 5.6 montre par exemple la fonction de corrélation $\xi(r_p, \pi)$ mesurée pour ces deux larges classes, avec un lissage gaussien réalisé sur une échelle de $3 h^{-1} \text{Mpc}$. J'ai ensuite calculé la projection $w_p(r_p)$ (à partir des fonctions $\xi(r_p, \pi)$ non lissées) pour chaque type de galaxies, pour chaque intervalle en *redshift*.

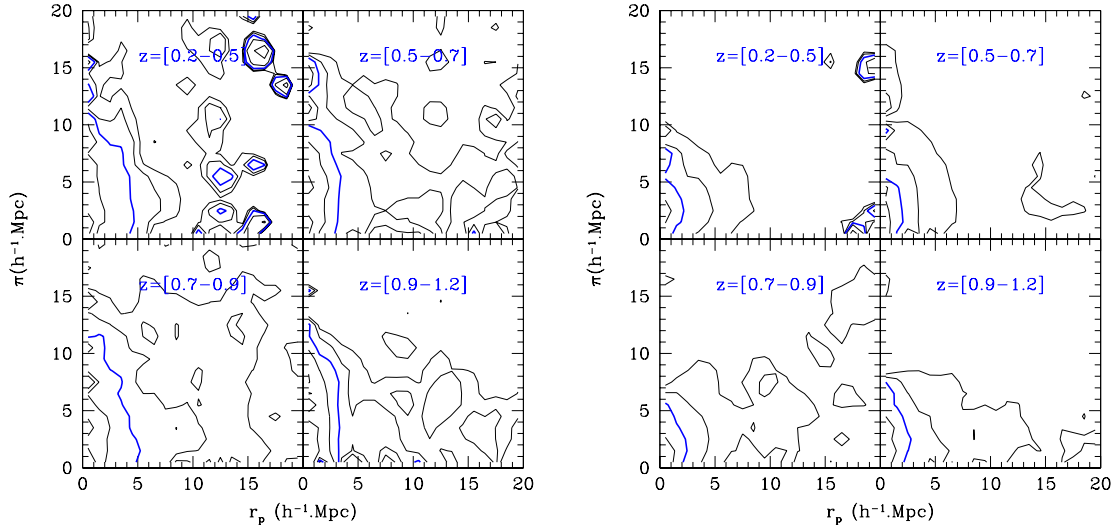


FIG. 5.6 – Fonctions $\xi(r_p, \pi)$ des échantillons spectroscopiques du VVDS-02h incluant les types 1+2 (à gauche) et 3+4 (à droite) dans quatre intervalles en *redshifts* consécutifs. Un lissage gaussien a été réalisé sur une échelle de $3 h^{-1} \text{Mpc}$. On a représenté les niveaux pour $\xi(r_p, \pi) = 0.2, 0.5, 1$ (en gras), 2 et 5

En moyenne, sur tout l'intervalle en *redshift* $[0.2-1.2]$, les galaxies de type 1+2 sont 1.8 fois plus agrégées que les galaxies de type 4. Ce résultat obtenu sur une grande plage de *redshifts* prolonge ceux obtenus dans l'Univers local (Loveday *et al.*, 1999; Norberg *et al.*, 2002; Zehavi *et al.*, 2002, 2005a) et à *redshifts* intermédiaires (Phleps *et al.*, 2005) présentés dans la partie 1.4.2.3.

Les galaxies précoces observées dans le VVDS à $z \sim 1$ ont une magnitude absolue comprise entre -22 et $-19 + 5 \log(h)$. Les galaxies observées dans l'Univers local par le 2dFGRS de luminosité similaire présentent une longueur de corrélation de $r_0 = 6$ à $7.5 h^{-1} \text{Mpc}$, croissante avec la luminosité (figure 1.19). L'amplitude de l'agrégation n'a donc pas évolué de plus de 70% entre *redshift* $z \sim 1$ et maintenant. Ce résultat semble

indiquer que les galaxies de type précoce étaient déjà en place au sein des structures de galaxies. Si l'on fait l'hypothèse que ces galaxies elliptiques ont simplement suivi l'évolution des structures sous l'effet de la gravitation, alors l'amplitude de la fonction de corrélation spatiale n'aurait pas évolué de plus de $\sigma_8(z=1)/\sigma_8(z=0) \simeq 0.6$ ce qui est de l'ordre de grandeur de ce que l'on observe.

La tendance observée de la longueur de corrélation des galaxies de type tardif semble indiquer que la majeure partie de la formation stellaire à $z \sim 1$ se déroule dans des galaxies plus fortement agrégées que ce que l'on observe localement. En prenant en compte l'évolution de la luminosité (Zucca *et al.*, 2005), les galaxies tardives du VVDS à $z \sim 1$ sont comparables à celles des sondages 2dFGRS, SDSS, Stromlo-APM dont les longueurs de corrélation sont de $r_0 \simeq 2.9-4.2 \text{ h}^{-1} \text{ Mpc}$ (Loveday *et al.*, 1999; Norberg *et al.*, 2002; Zehavi *et al.*, 2002). Une interprétation possible consiste à dire que la formation stellaire se déplace des régions les plus denses vers les régions les moins denses à mesure que le *redshift* diminue.

Pour compléter cette étude par type, j'ai séparé l'ensemble des galaxies en deux populations suivant leur couleur. En effet, ainsi que je l'ai dit dans la partie 1.5.1.5, il est observé dans l'Univers local une bimodalité en couleur rouge/bleue (Strateva *et al.*, 2001). Cette division a également été observée à grand *redshifts* (van Dokkum *et al.*, 2000; Bell *et al.*, 2004; Giallongo *et al.*, 2005; Willmer *et al.*, 2005). Nous avons appliqué la relation de Giallongo *et al.* (2005) à nos données. La figure 5.7 met en évidence les deux populations rouge et bleue, entre $z=0.2$ et $z=1.2$, réparties autour de la relation

$$\langle U - V \rangle = -0.098(M_B + 20) + 1.64 \quad (5.2)$$

valable selon Giallongo *et al.* (2005) pour $0.4 \leq z \leq 1$ mais ici étendue jusqu'à $z=1.2$ pour les galaxies rouges et $z=2.0$ pour les galaxies bleues.

Après avoir divisé les galaxies en échantillon rouge/bleu, on calcule la fonction de corrélation $\xi(r_p, \pi)$ reproduite figure 5.8 (après lissage). Les projections $w_p(r_p)$ sont ensuite calculées et les paramètres r_0 et γ déterminés avec leurs erreurs associées (figure 5.9). L'évolution de la longueur de corrélation de la population bleue est quasiment identique à celle observée pour les galaxies de type 3&4. La valeur de r_0 pour la classe rouge évolue très peu entre $z \sim 1$ et $z \sim 0.3$, ce qui est la principale différence entre l'analyse par type spectral ou par couleur de l'agrégation des galaxies du VVDS.

Comme observé dans l'Univers local (Brown *et al.*, 2000; Zehavi *et al.*, 2002, 2005b, voir partie 1.5.1.5), l'amplitude de la fonction de corrélation de la population rouge est plus importante que celle de la population bleue, de même, sa pente est plus forte. Cette ségrégation est clairement marquée dès $z \sim 1$. Les galaxies rouges, constituées d'étoiles vieilles sont donc déjà situées dans les régions denses, alors que les galaxies bleues, constituées d'étoiles jeunes sont d'avantage diluées dans le champ. Le lien avec la classification spectrale est évidente. Malgré une dichotomie différente, les deux définitions sont fortement corrélées. La classe rouge est en effet constituée en moyenne de 89% de galaxies de type 1 ou 2 jusqu'à $z=1.2$, alors qu'en moyenne, 91% des galaxies bleues sont de type 3 ou 4 jusqu'à $z=2$.

C'est la première fois qu'une analyse des propriétés d'agrégation des galaxies selon leur type spectral et leur couleur est possible sur une si grande plage en *redshift*. Les résultats sont soumis pour publication à la revue *Astronomy & Astrophysics* (section 5.2.4).

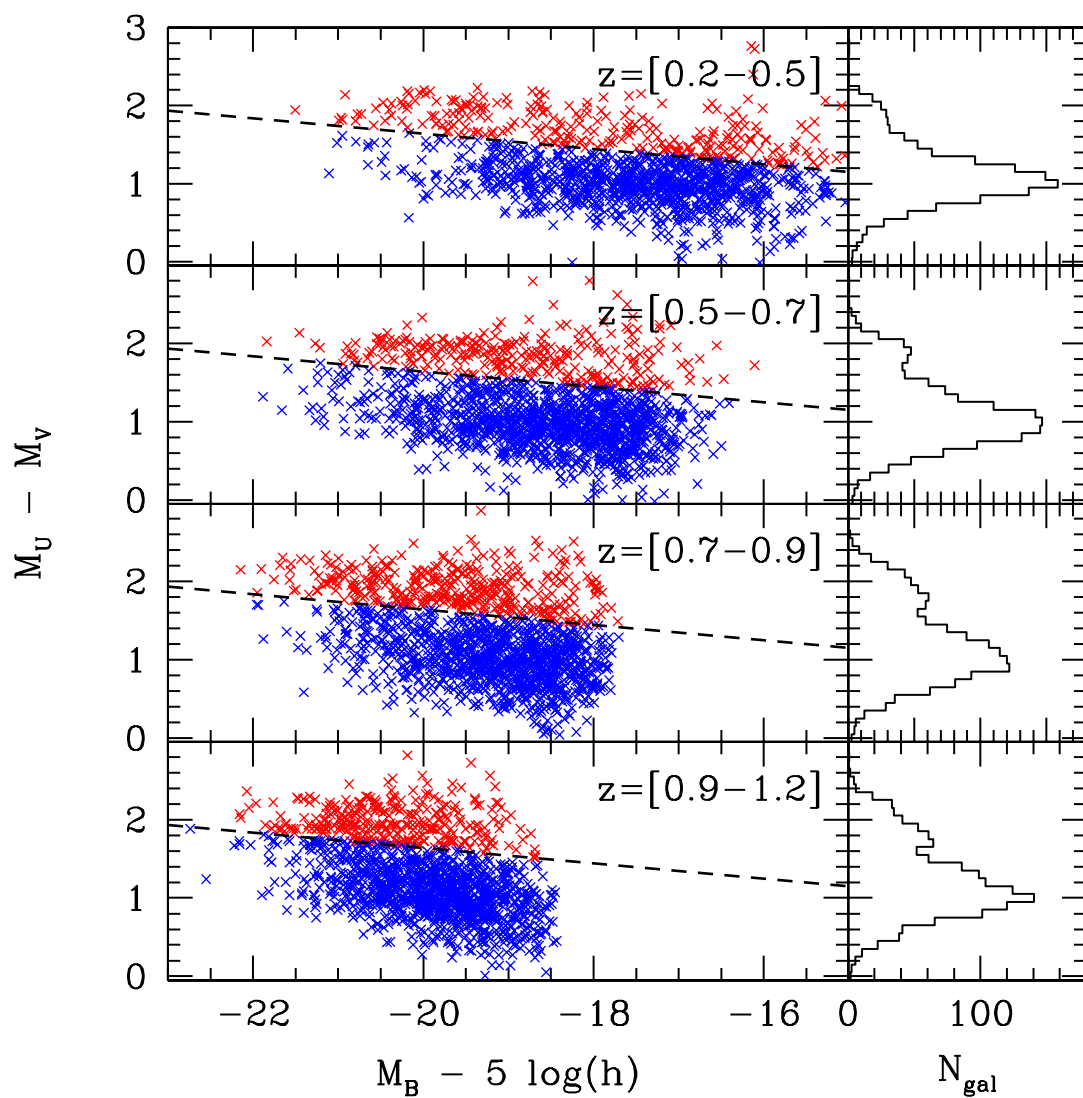


FIG. 5.7 – Division couleur-magnitude dans le champ VVDS-02h selon la relation de [Giallongo et al. \(2005\)](#) (ligne pointillé noire), pour quatre intervalles de *redshifts*. À droite, les distributions en couleur (U-V) au repos.

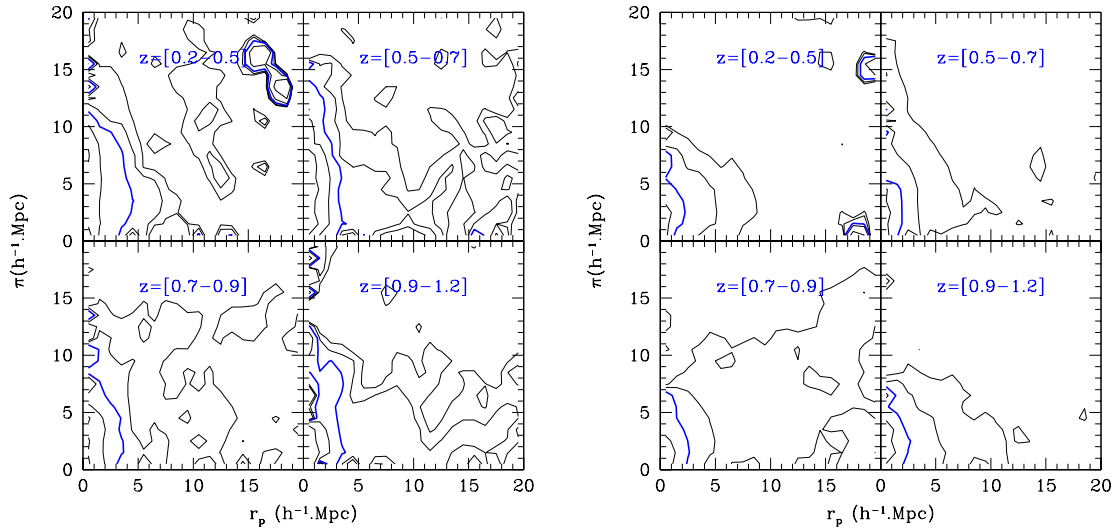


FIG. 5.8 – Fonctions $\xi(r_p, \pi)$ des populations de galaxies rouges (à gauche) et bleues (à droite) dans le champ VVDS-02h mesurées dans quatre intervalles en *redshifts* consécutifs. Un lissage gaussien a été réalisé sur une échelle de $3 \text{ h}^{-1} \text{ Mpc}$. On a représenté les niveaux pour $\xi(r_p, \pi)=0.2, 0.5, 1$ (en gras), 2 et 5

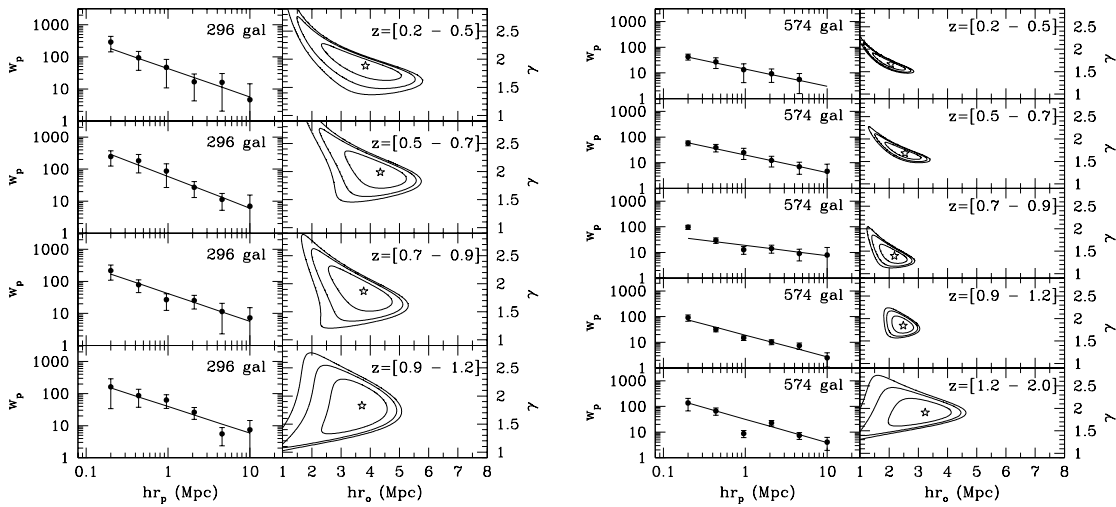


FIG. 5.9 – Fonctions de corrélation $w_p(r_p)$ calculées pour les galaxies rouges (à gauche) et bleues (à droite) en fonction du *redshift*, avec les contours d'erreurs à 1, 2 et 3σ associés aux valeurs de r_0 et γ .

5.2.4 Article III : *The VIMOS VLT Deep Survey - the evolution of galaxy clustering per spectral type to $z \sim 1.5$*

Cet article a été soumis pour publication dans la revue européenne *Astronomy & Astrophysics* par Meneux, B., Le Fèvre, O., Guzzo, L., Pollo, A. *et al.* (astro-ph/0511656). Il présente la mesure et l'évolution de la fonction de corrélation spatiale des galaxies du sondage VVDS selon leur type spectral et leur couleur, jusqu'à $z \sim 2$.

The VIMOS-VLT Deep Survey

The evolution of galaxy clustering per spectral type to $z \simeq 1.5$ *

B. Meneux¹, O. Le Fèvre¹, L. Guzzo², A. Pollo^{1,2}, A. Cappi³, O. Ilbert⁴, A. Iovino⁵, C. Marinoni^{5,6}, H.J. McCracken^{7,8}, D. Bottini⁹, B. Garilli⁹, V. Le Brun¹, D. Maccagni⁹, J.P. Picat¹⁰, R. Scaramella¹¹, M. Scodeggio⁹, L. Tresse¹, G. Vettolani¹², A. Zanichelli¹², C. Adami¹, S. Arnouts¹, M. Arnaboldi¹³, S. Bardelli³, M. Bolzonella⁴, S. Charlot^{7,14}, P. Cilieggi³, T. Contini¹⁰, S. Foucaud¹⁵, P. Franzetti⁹, I. Gavaud^{10,16}, B. Marano⁴, A. Mazure¹, R. Merighi³, S. Paltani^{17,18}, R. Pello¹⁰, L. Pozzetti³, M. Radovich¹³, G. Zamorani³, E. Zucca³, M. Bondi¹², A. Bongiorno⁴, G. Busarello¹³, O. Cucciati^{2,14}, L. Gregorini¹², F. Lamareille¹⁰, G. Mathez¹⁰, Y. Mellier^{7,8}, P. Merluzzi¹³, V. Ripepi¹³, D. Rizzo¹⁰

¹ Laboratoire d'Astrophysique de Marseille, UMR 6110 CNRS-Université de Provence, BP8, 13376 Marseille Cedex 12, France

² INAF-Osservatorio Astronomico di Brera - Via Bianchi 46, I-23807, Merate, Italy

³ INAF-Osservatorio Astronomico di Bologna - Via Ranzani, 1, I-40127, Bologna, Italy

⁴ Università di Bologna, Dipartimento di Astronomia - Via Ranzani, 1, I-40127, Bologna, Italy

⁵ INAF-Osservatorio Astronomico di Brera - Via Brera 28, Milan, Italy

⁶ Centre de Physique Theorique, UMR 6207 CNRS-Universite de Provence, case 907 F-13288 Marseille, France

⁷ Institut d'Astrophysique de Paris, Université Pierre et Marie Curie, UMR 7095, 98 bis Bvd Arago, 75014, Paris, France

⁸ Observatoire de Paris, LERMA, 61 Avenue de l'Observatoire, 75014 Paris, France

⁹ INAF - IASF Milano, via Bassini 15, I-20133 Milano, Italy

¹⁰ Laboratoire d'Astrophysique de l'Observatoire Midi-Pyrénées (UMR 5572) - 14, avenue E. Belin, F31400 Toulouse, France

¹¹ INAF-Osservatorio Astronomico di Roma - Via di Frascati 33, I-00040, Monte Porzio Catone, Italy

¹² IRA-INAf - Via Gobetti, 101, I-40129, Bologna, Italy

¹³ INAF-Osservatorio Astronomico di Capodimonte - Via Moiariello 16, I-80131, Napoli, Italy

¹⁴ Max Planck Institut für Astrophysik, 85741, Garching, Germany

¹⁵ School of Physics & Astronomy, University of Nottingham, Nottingham NG7 2RD, England

¹⁶ European Southern Observatory, Karl-Schwarzschild-Strasse 2, D-85748, Garching bei München, Germany

¹⁷ Integral Science Data Centre, ch. d'Écogia 16, CH-1290 Versoix, Switzerland

¹⁸ Geneva Observatory, ch. des Maillettes 51, CH-1290 Sauverny, Switzerland

¹⁹ Università di Milano-Bicocca, Dipartimento di Fisica - Piazza delle Scienze, 3, I-20126 Milano, Italy
email: baptiste.meneux@oamp.fr

Received ... 2005 ; accepted ... 2005

Abstract. We measure the evolution of clustering for galaxies with different spectral types from 6495 galaxies with $17.5 \leq I_{AB} \leq 24$ and measured spectroscopic redshift in the first epoch VIMOS-VLT Deep Survey (VVDS). We classify our sample into four classes, based on the fit of well-defined galaxy spectral energy distributions on observed multi-color data. We measure the projected function $w_p(r_p)$ and estimate the best-fit parameters for a power-law real-space correlation function $\xi(r) = (r/r_0)^{-\gamma}$. We find the clustering of early-spectral-type galaxies to be markedly stronger than that of late-type galaxies at all redshifts up to $z \simeq 1.2$. At $z \sim 0.8$, early-type galaxies display a correlation length $r_0 = 4.8 \pm 0.3 \text{ h}^{-1}\text{Mpc}$, while late types have $r_0 = 2.5 \pm 0.2 \text{ h}^{-1}\text{Mpc}$. For the latest class of star-forming blue galaxies, we are able to push our clustering measurement to an effective redshift $z \sim 1.4$, for luminous galaxies ($M_B(AB) \simeq -21$). The clustering of these objects increases up to $r_0 = 3.42 \pm 0.4 \text{ h}^{-1}\text{Mpc}$ for $z = [1.2, 2.0]$. The relative bias between early- and late-type galaxies within our magnitude-limited survey remains approximately constant with $b \sim 1.7 - 1.8$ from $z \simeq 0.2$ up to $z \simeq 1$, with indications for a decrease at $z > 1.2$, due to the growth in clustering of the star-forming population. We find similar results when splitting the sample into “red” and “blue” galaxies using the observed color bi-modality. When compared to the expected linear growth of mass fluctuations, a natural interpretation of these observations is that: (a) the assembly of massive early type galaxies is already mostly complete in the densest dark matter halos at $z \simeq 1$; (b) luminous late-type galaxies are located in higher-density, more clustered regions of the Universe at $z \simeq 1.5$ than at present, indicating that star formation activity is progressively increasing, going back in time, in the higher-density peaks that today are mostly dominated by old galaxies.

Key words. surveys - galaxies: evolution - cosmology: large scale structure of Universe

1. Introduction

The measurement of the evolution of the clustering of galaxies is a key element to understand the evolution of the Universe and the formation of large-scale structures. In the current paradigm, the structures we observe today in the large-scale distribution of galaxies are the result of the gravitational amplification of primordial overdensities in the dark matter density field. Smaller scales are supposed to collapse first, with virialized structures as galaxies, groups and clusters subsequently built up via multiple merging, at a rate of growth essentially governed by the value of the mean matter density Ω_M . In this hierarchical formation scenario, the evolution of the clustering of dark matter halos is now well understood both analytically (Mo & White 1996; Sheth & Tormen 1999) and from N-body simulations (Jenkins et al. 1998; Kauffmann et al. 1999a; Springel et al. 2005).

The visible component of galaxies is expected to assemble via the collapse of baryonic matter within the dark matter halos. A naive picture would let to conclude that galaxies follow the underlying dark matter density field. But the formation and evolution of galaxies is driven by physical processes such as cooling, star formation and feedback, which are not easily linked to the underlying dark matter halos (Kauffmann et al. 1999a). Rather, a general prediction of most current models is that galaxy formation is favoured in higher-density regions. Objects forming in regions of larger-than-average density collapse first and, looking back in time, are found to inhabit higher and higher density peaks (i.e. they are more *biased*). Recent direct measurements of the evolution of galaxy bias up to a redshift $z \sim 1.5$ (Marinoni et al. 2005) provide supporting evidence to this biased picture of galaxy formation.

Although being well known since more than 25 years, the origin of the relationship between galaxy morphological (or almost equivalently, spectral) type and local structure is still not fully understood (e.g. Dressler 1980; Smith et al. 2005; Postman et al. 2005). Whether the higher fraction of early-type galaxies in high-density regions is established *ab initio* (the so-called *nature* hypothesis), or rather is produced by environmental effects during the life of the galaxy (*nurture*) is still debated. A practical effect of the morphology-density relation is that early- and late-type galaxies trace differently the density field, with the former class being more clustered (i.e. more *biased*) than the latter. In addition, more luminous galaxies (which are

preferentially early-types) are more clustered than less luminous ones, with the two parameters – luminosity and spectral type – playing together in a subtle way (Iovino et al. 1993; Madgwick et al. 2003; Norberg et al. 2002; Zehavi et al. 2002; Loveday Tresse & Maddox 1999; Guzzo et al. 1997; Benoist et al. 1996). Since the amount of biasing should be related to the mass and formation history of a galaxy, one therefore expects the evolution of the clustering of galaxies with different morphological types to contain important clues on the assembly of the baryonic mass and the origin of galaxy types.

Large deep surveys show consistently that the clustering of the overall galaxy population evolves slowly with time (Coil et al. 2004; Le Fèvre et al. 2005a). Measurements from the VIMOS-VLT Deep Survey (VVDS, Le Fèvre et al. 2005a), indicate that the clustering amplitude of L^* galaxies increases by a factor ~ 2.4 from $z = 1.5$ to the present epoch. However, little is known about the clustering evolution of different galaxy types above $z \sim 0.3$. In the CFRS sample, no significant difference was observed between the clustering of red and blue galaxies up to $z \sim 1$ (Le Fèvre et al. 1996), although the sample was too small to reach firm conclusions. Red galaxies at $z \approx 0.6$ in the CNOC2 survey were found to have a clustering amplitude ~ 5 times higher than blue galaxies (Carlberg et al. 1997). Results from the CADIS survey using photometric redshifts indicate that up to $z \sim 1$ early spectral type galaxies are more clustered, with a correlation amplitude 2.1 times larger than late type at $z=[0.75-1.07]$ (Phleps & Meisenheimer 2003). A more recent and accurate estimate at $\langle z \rangle = 0.6$ (Phleps et al. 2005) confirms the same trend among red and blue galaxies with $r_0^{red} = 5.39^{+0.30}_{-0.28} h^{-1} \text{Mpc}$ and $r_0^{blue} = 3.64^{+0.25}_{-0.24} h^{-1} \text{Mpc}$. At even higher redshifts, the measurement of type is substituted by the selection of specific sub-populations, e.g. Daddi et al. (2003) measure a strong clustering of extremely red (J-K) galaxies at $2 \leq z \leq 4$, arguing that these may represent the progenitors of massive elliptical galaxies at lower redshifts.

In this paper we attempt to accurately quantify the contribution of different galaxy types to the observed evolution of clustering using the VVDS first epoch survey. The accuracy of spectroscopic redshifts is used to project the correlation function $\xi(r_p, \pi)$ in narrow velocity intervals and to measure the correlation length r_0 and the slope γ of the real space correlation function as a function of redshift up to $z \sim 1.2$.

We describe the sample and our galaxy classification in Section 2, and the method used to compute the real space correlation function in Section 3. The clustering measurements are presented in Section 4, and then discussed in Section 5. We use a cosmology with $\Omega_M = 0.3$, $\Omega_\Lambda = 0.7$, throughout the paper. All distances, given in comoving unit, and absolute magnitudes are computed with $h = 1$ (which denote $H_0/100 \text{ km s}^{-1} \text{Mpc}^{-1}$).

2. VVDS first epoch data

2.1. The sample

We are using the VIMOS-VLT Deep Survey first epoch data described in Le Fèvre et al. (2005). We concentrate on the VVDS-02h field, and we keep all galaxies with the best redshifts avail-

Send offprint requests to: baptiste.meneux@oamp.fr

* based on data obtained with the European Southern Observatory Very Large Telescope, Paranal, Chile, program 070.A-9007(A), and on data obtained at the Canada-France-Hawaii Telescope, operated by the CNRS of France, CNRC in Canada and the University of Hawaii and observations obtained with MegaPrime/MegaCam, a joint project of CFHT and CEA/DAPNIA, at the Canada-France-Hawaii Telescope (CFHT) which is operated by the National Research Council (NRC) of Canada, the Institut National des Science de l'Univers of the Centre National de la Recherche Scientifique (CNRS) of France, and the University of Hawaii. This work is based in part on data products produced at TERAPIX and the Canadian Astronomy Data Centre as part of the Canada-France-Hawaii Telescope Legacy Survey, a collaborative project of NRC and CNRS.

Table 1. The four VVDS galaxy types

type		number of objects	indicated rest-frame colour
1	E/S0	645	$1.3 < (B_{AB} - I_{AB})$
2	early spiral	1004	$0.95 < (B_{AB} - I_{AB}) < 1.3$
3	late spiral	2104	$0.68 < (B_{AB} - I_{AB}) < 0.95$
4	irregular starburst	2742	$(B_{AB} - I_{AB}) < 0.68$

able, with quality flags 2 to 9 (confidence level greater than 80% in the redshift measurement). AGN are excluded from the sample. The analysed sample in this paper includes 6495 galaxies with spectroscopic redshifts in an area of 1750 arcmin². The accuracy of the redshift measurement is $\sim 275 \text{ km.s}^{-1}$. A complete presentation of the sample is given in Le Fèvre et al. (2005).

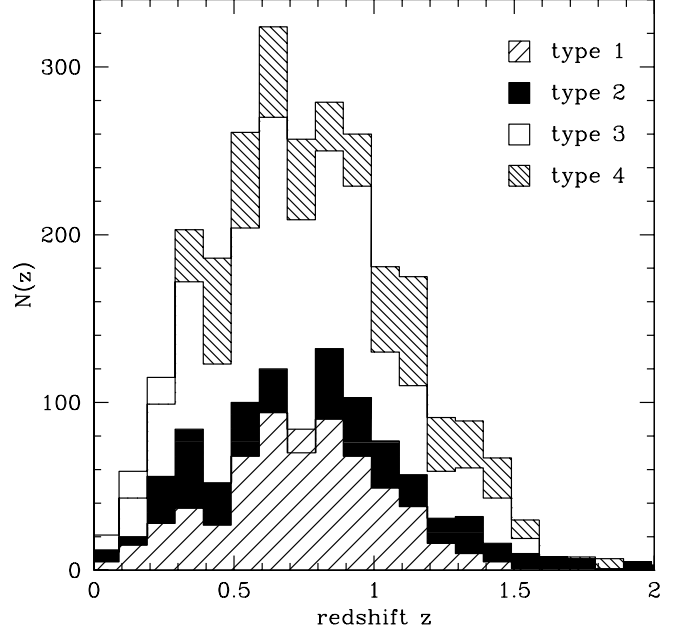
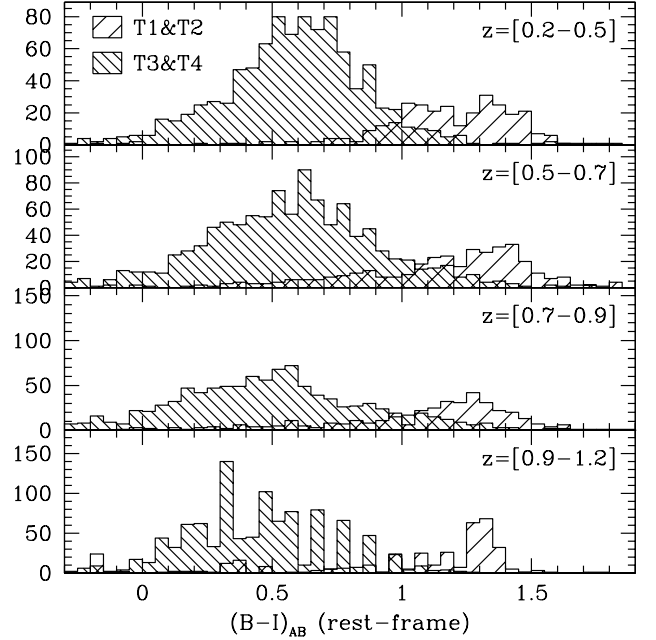
2.2. Galaxy classification

We used the galaxy classification introduced by Zucca et al. (2005) to study the evolution of the luminosity function per galaxy type from the first epoch VVDS data. The classification has been obtained from the B to I band multi-wavelength dataset available in the VVDS-02h field (Le Fèvre et al. 2004; McCracken et al. 2003), and the spectroscopic redshifts (Le Fèvre et al. 2005). The U band, which does not cover all the field (Radovich et al. 2003) was not used in order to get a homogeneous classification on the sky. Rest-frame colors have been matched to the empirical set of spectral energy distribution (SED) templates described in Arnouts et al. (1999), made of four observed spectra (Coleman Wu & Weedman 1980, CWW hereafter) and two starburst SEDs computed with GISSSEL (Bruzual & Charlot 1993).

The original spectroscopic sample of 6495 galaxies was divided into four spectral classes from early-type (type 1) to irregular / star-forming galaxies (type 4) based on the CWW templates. The later type also includes galaxies with SEDs well represented by starburst templates. The properties of these four classes of galaxies are summarized in Table 1 and in the redshift distributions shown in Figure 1. As a result of the VVDS I-band selection, we are increasingly loosing red early-type galaxies for redshifts $z > \sim 1.2$ (figure 2) because galaxies are selected from their rest frame UV flux at these redshifts, but late type galaxies continue to be identified up to the highest redshifts discussed in this paper.

As described in Zucca et al. (2005), the evolution of the global luminosity function in the VVDS is strongly driven by type 4 galaxies, which evolve by $\simeq 2$ magnitudes from $z \sim 1.5$ and present a steep faint end luminosity function slope. Conversely, the luminosity function of early type galaxies remains stable over this redshift range. The relative distribution of absolute magnitudes per type presented in Figure 3 is the result of this evolution.

In summary, the type classification used in this paper is a spectral classification based on a set of spectral templates covering from early type galaxies to late and star-forming galaxies.

**Fig. 1.** Redshift distribution of the four galaxy types**Fig. 2.** Rest-frame color $(B - I)_{AB}$ for early-type galaxies (type 1&2) and late-type galaxies (type 3&4).

Although it is tempting to associate spectral type to morphological type, and a close relationship exists between them, the spectral classification used in this paper is more closely related to a color classification from “red” to “blue” galaxies.

3. Measurement and associated errors: method

The method to recover the correlation length r_0 and the slope γ of the real space correlation function is well detailed in the pa-

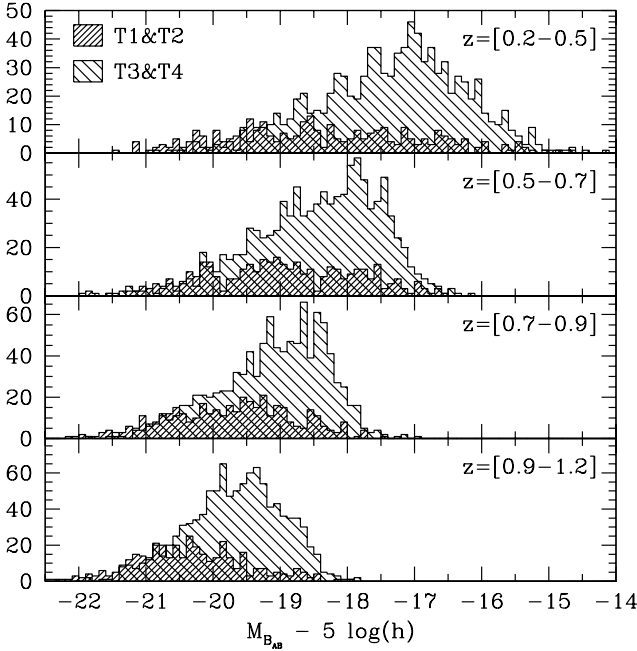


Fig. 3. B-band luminosity distribution $M_B(AB)$ for early-type galaxies (types 1&2) and late-type galaxies (types 3&4)

per of Pollo et al. (2005). We are focussing here on the most critical aspects of the method when applied to compute the clustering properties as a function of galaxy type.

3.1. The real space correlation functions parameters

We computed the bi-dimensional two-point correlation function $\xi(r_p, \pi)$ for each galaxy type, in different redshift bins chosen to maximize the number of objects hence the signal to noise ratio of the clustering measurement. This was estimated using the Landy & Szalay (1993) estimator

$$\xi(r_p, \pi) = \frac{N_R(N_R - 1)}{N_G(N_G - 1)} \frac{GG(r_p, \pi)}{RR(r_p, \pi)} - 2 \frac{N_R - 1}{N_G} \frac{GR(r_p, \pi)}{RR(r_p, \pi)} + 1 \quad (1)$$

where N_G is the mean galaxy density (or, equivalently, the total number of objects) in the survey; N_R is the mean density of a catalog of random points distributed within the same survey volume; $GG(r_p, \pi)$ is the number of independent galaxy-galaxy pairs with separation between r_p and $r_p + dr_p$ perpendicular to the line-of-sight and between π and $\pi + d\pi$ along the line of sight; $RR(r_p, \pi)$ is the number of independent random-random pairs within the same interval of separations and $GR(r_p, \pi)$ represents the number of galaxy-random cross pairs.

The real-space correlation function $\xi(r)$ represents the excess probability for a given pair of galaxies to be observed with a separation r (Peebles 1980). It can be derived from $\xi(r_p, \pi)$ using the formalism of Davis & Peebles (1983), computing $w_p(r_p)$, the projection of $\xi(r_p, \pi)$ along the line of sight.

$$w_p(r_p) \equiv 2 \int_0^\infty \xi(r_p, \pi) d\pi = 2 \int_0^\infty \xi[(r_p^2 + \pi^2)^{1/2}] d\pi \quad (2)$$

In practice, the upper integration limit must be chosen finite, as to include the real signal, without adding extra noise which

is dominant above a certain π . Based on direct tests with the GALICS mock samples (Blaizot et al. 2005), we choose a value $\pi_{max} = 20 \text{ h}^{-1} \text{Mpc}$. This is high enough to sum the real signal, without adding too much noise. It is compatible with the VVDS velocity measurement error of $\sim 275 \text{ km.s}^{-1}$. Assuming a power-law form for the real-space correlation function $\xi(r) = (r/r_0)^\gamma$, the integral in equation 2 can be solved analytically in terms of Gamma functions, yielding a $w_p(r_p)$ which is itself a power law. The fitting method of the measured $w_p(r_p)$, detailed in Pollo et al. (2005), provides the values of the correlation length r_0 and the slope γ with their associated errors. A specific treatment have been applied to compute the clustering properties per galaxy types, as described in Sections 3.2 and 3.3.

3.2. Biases and weights for each spectral class

In order to correct for the observational biases introduced by the VVDS observing strategy in the VVDS-02h field, we used the same correction scheme explained in Pollo et al. (2005). The main biases are the incompleteness in redshift measurement and the complex geometry of the field. This results in a spatial sampling that changes as a function of position in the field. To account for galaxies with unknown redshift (either unobserved or discarded because of a low redshift quality flag), the observed galaxies are given a weight.

The goal of this weight is to recover the true number of galaxy pairs at a given angular scale.

The computation of this local weight takes into account the local number of galaxies (with and without a redshift) around a given galaxy in the spectroscopic sample. The parent photometric sample used for this computation must have the same properties as the spectroscopic one. For instance, the weight computed for the early type galaxies spectroscopic sample must make reference only to the early type galaxies in the underlying complete photometric sample, without including other galaxy types. In this approach, the parent photometric catalogue used to compute the weight must therefore only contain galaxies with the same properties than those of the spectroscopic sample.

Photometric redshifts of all the galaxies in the VVDS-02h have been derived from BVRI photometry (Le Fèvre et al. 2004; McCracken et al. 2003), and ugriz bands from the CFHT Legacy Survey. Details are presented in Ilbert et al. (in prep.). The photometric redshifts and rest-frame magnitudes were then used to give a spectral type to all these galaxies, according to the method explained in Section 2.2. The complete parent photometric catalogue was then divided in four catalogues, one for each type. Note that the photometric redshifts are only used to define a spectral type to each galaxy in the whole photometric catalog in order to compute the weights associated to each galaxy type, in addition to the spectroscopically measured galaxies. Photometric redshifts are not directly used here in the computation of the correlation function.

3.3. Error measurement

Even if our spectroscopic sample is the largest to date at this depth, and despite the angular size of the VVDS-02h (1750 arcmin²), the uncertainties associated to the measurement of the power-law parameters of the real-space correlation function $\xi(r)$ are largely dominated by cosmic variance.

To assess this, we have constructed 50 VVDS-02h mock surveys from the GalICS simulation (Blaizot et al. 2005). GalICS is a semi-analytic model for galaxy evolution post-processed to a large cosmological N-body simulation. These mock surveys are quasi-independent and are built applying all the VVDS-02h observing biases: the complex geometry of the field, the spectroscopic selection function, the redshift measurement accuracy, the incompleteness (see Pollo et al. (2005) for more information). We use the clustering variance in the mock samples to quantify the cosmic variance expected on the correlation length r_0 and the slope γ of the real-space correlation function.

Our method to assign a spectral type to a given galaxy does not work so well with the GalICS simulations. We were not able to recover a realistic redshift distribution for the four spectral types using the CFH12k BVRI apparent magnitudes produced by GalICS. For example, there were no Type 1 (elliptical) galaxies beyond $z \sim 0.5$ in most of the 50 GalICS cones. Conversely, the galaxy type classification given in for GalICS data do not match ours. This “Morphological” type classification is based on rest-frame B-band (bulge+burst)-to-disc luminosity ratio. This provides a huge “spiral” class (on average $\sim 90\%$ of the full dataset), and two equivalent (in number of objects) “elliptical” and “lenticular” classes. These two last classes could match our type 1 class. But then, this is quite difficult to divide the huge “spiral” class in our 3 later types.

Computing error measurements from a bootstrap resampling method slightly under-estimates the real variance: if we consider the full VVDS-02h data set (Le Fèvre et al. 2005a), the relative errors computed from GalICS are $\sim 1.5 - 2$ times greater than error computed from bootstrap resampling, we expect similar factor when dividing our sample per spectral type.

Therefore the approach that has been chosen to compute measurement errors consists in taking randomly, in each of the 50 mock VVDS-02h surveys, a proportion of galaxies equal to that found in the VVDS-02h spectroscopic sample for a given spectral class. Even if the clustering measured is the one of the “full” population with higher uncertainties, the statistic of each “type” is representative of the real one. The variance from field to field is artificially increased by a broader range of clustering properties in the simulated data, and therefore, this method is very conservative in the way that it is expected to produce errors for the observed data larger than in reality.

4. Clustering evolution for each galaxy type

4.1. Measurement of r_0 and γ

We have computed the correlation function $\xi(r_p, \pi)$ and its projection $w_p(r_p)$ in the VVDS-02h field, for each galaxy type of our classification, in increasing redshift slices. In order to improve the clustering signal, and depending on the redshift bin

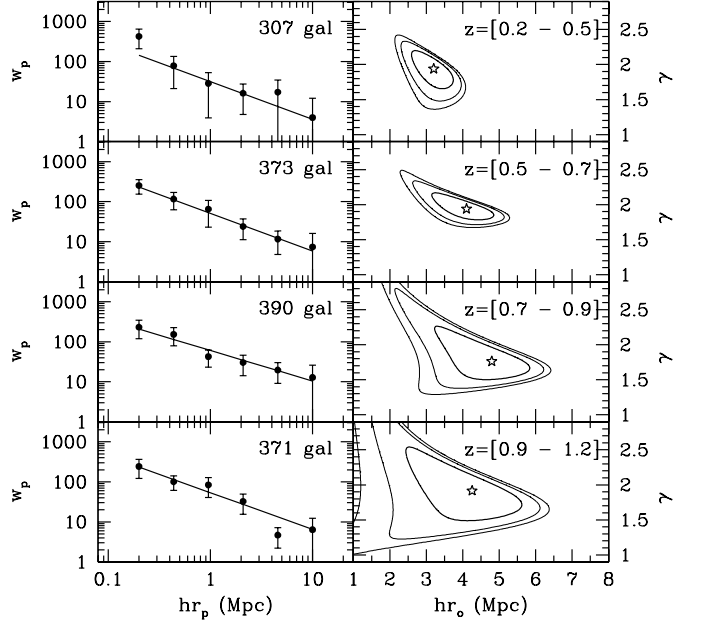


Fig. 4. (left) Correlation function $w_p(r_p)$ as a function of redshift for early-type galaxies with types 1 and 2 (right). Error contours (68%, 90% and 95% likelihood levels) associated to the measurement of the correlation length r_0 and slope γ of the correlation function.

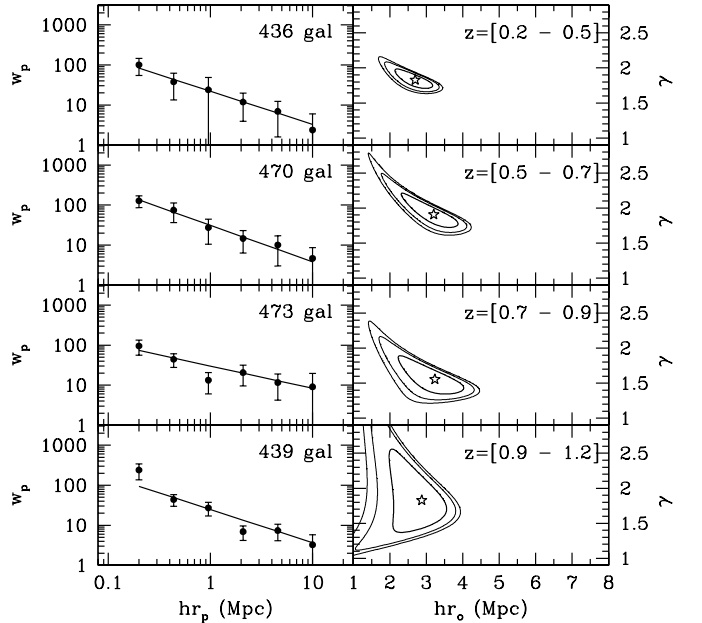


Fig. 5. Same as Figure 4 for galaxies with type 3.

considered, we merged the two earliest-type galaxy samples $T1$ and $T2$ and the two latest types $T3$ and $T4$. Merging types 1 and 2 would clearly dilute the clustering signal of type 1 galaxies if they are more strongly clustered than type 2. The measurements of $w_p(r_p)$ are presented in Figures 4 to 7.

The projection $w_p(r_p)$ has been fitted by a power-law on scales $0.1 \leq r_p \leq 10 h^{-1} \text{Mpc}$. The values of the correlation

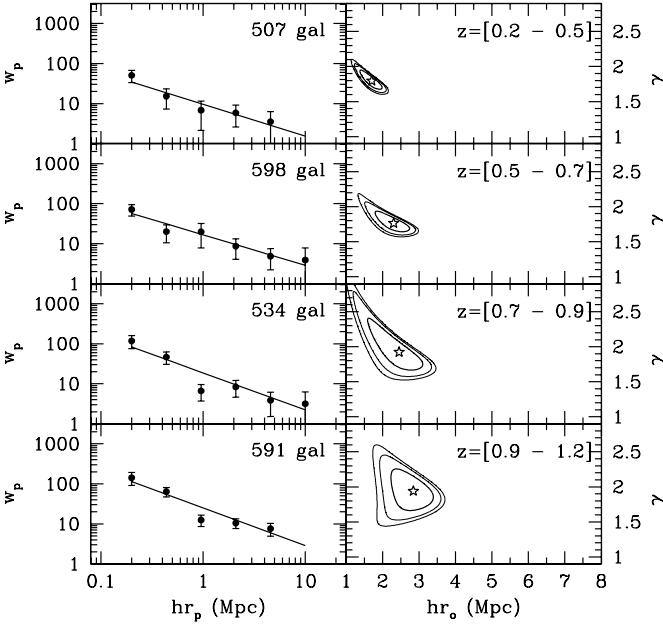


Fig. 6. Same as Figure 4 for galaxies with type 4.

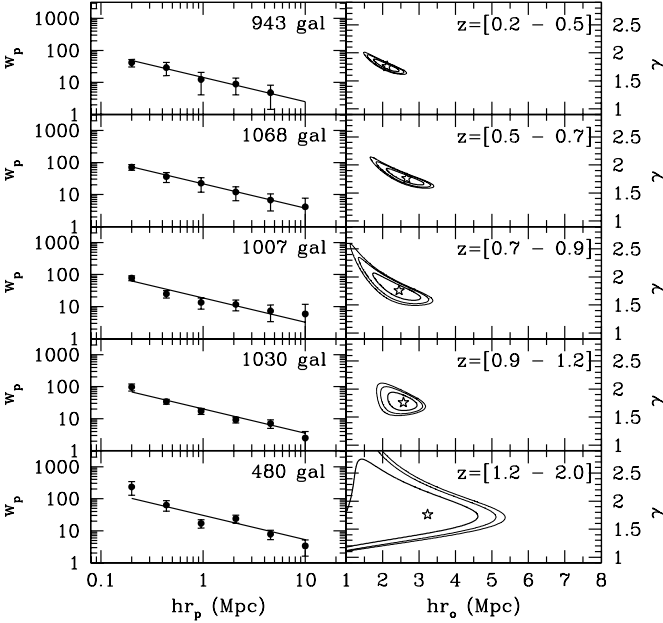


Fig. 7. Same as Figure 4 for galaxies with type 3 and 4.

length r_0 and the slope γ of the real-space correlation function are summarized in table 2. A clear difference in clustering strength is observed between the 3 classes of galaxies. We find that on average the clustering strength of early type galaxies is 1.8 times larger than the clustering strength of type 4 galaxies. This is observed over the full redshift range $0.2 \leq z \leq 1.2$ and this result is significant. The probability that the correlation lengths of these two populations are the same is indeed

only 2.3% for the redshift bin $z=[0.2,0.5]$, 2.7% for $z=[0.5,0.7]$, 2.2% for $z=[0.7,0.9]$, and 14.5% for $z=[0.9,1.2]$.

The evolution of the clustering length for the different classes of galaxies is presented in figure 8. An overall increase in the clustering of all galaxy types is observed up to $z \simeq 1$. In the range $z=[0.9,1.2]$, there is a slight indication for a weakening in the clustering of early type galaxies. The clustering length of late type / star-forming galaxies continues to rise up to redshift ~ 2 .

4.2. Relative bias and its evolution

Since it is not possible to measure directly the real-space correlation function of the mass, any attempt to measure the bias b from galaxy samples has either to assume the knowledge of the background cosmological parameters (*e.g.* Marinoni et al. 2005), or to obtain the mass variance from other observables like the Microwave Background (*e.g.* Lahav et al. 2002).

The simplest definition of the bias b that links linearly the mass density field ρ of the mass to the galaxy density field n is (for a fixed scale R over which ρ and n are measured)

$$\frac{\delta n}{\langle n \rangle} = b \frac{\delta \rho}{\langle \rho \rangle}, \quad (3)$$

where b is simply a number, independent on scale and on the local value of $\delta \rho / \langle \rho \rangle$ hypothesis. Similarly, but in a statistical rather than deterministic way, the bias can be defined in terms of *rms* values of the galaxy and density fields

$$\left(\frac{\delta n}{\langle n \rangle} \right)_{rms} = b \left(\frac{\delta \rho}{\langle \rho \rangle} \right)_{rms}. \quad (4)$$

With this definition, b is simply the ratio of the square root of the variance in the galaxy and mass density fields measured over a given scale R , i.e. $\sigma_R(\text{galaxies}) = b \sigma_R(\text{mass})$. In our case, we can measure directly the relative bias between two populations A and B, as

$$\frac{b_A}{b_B} = \frac{\sigma_8^A(r)}{\sigma_8^B(r)}, \quad (5)$$

where σ_8 can be obtained from the measured correlation function as

$$\sigma_R^2 = \frac{3}{R^3} \int_0^R r^2 \xi(r) dr, \quad (6)$$

that, for power-law $\xi(r)$, can be written as (Peebles 1980)

$$\sigma_8 = \sqrt{C_\gamma \left(\frac{r_0}{8 \text{ Mpc/h}} \right)^\gamma}, \quad (7)$$

with

$$C_\gamma = \frac{72}{(3-\gamma)(4-\gamma)(6-\gamma)2^\gamma}. \quad (8)$$

Figure 9 shows the relative bias between the populations T1+T2 and T4 at different redshifts, measured using eq. 7 in spheres with radius $r = 8 \text{ h}^{-1} \text{ Mpc}$. The average value is $b = 1.7 - 1.8$, consistent with a constant value with redshift up to

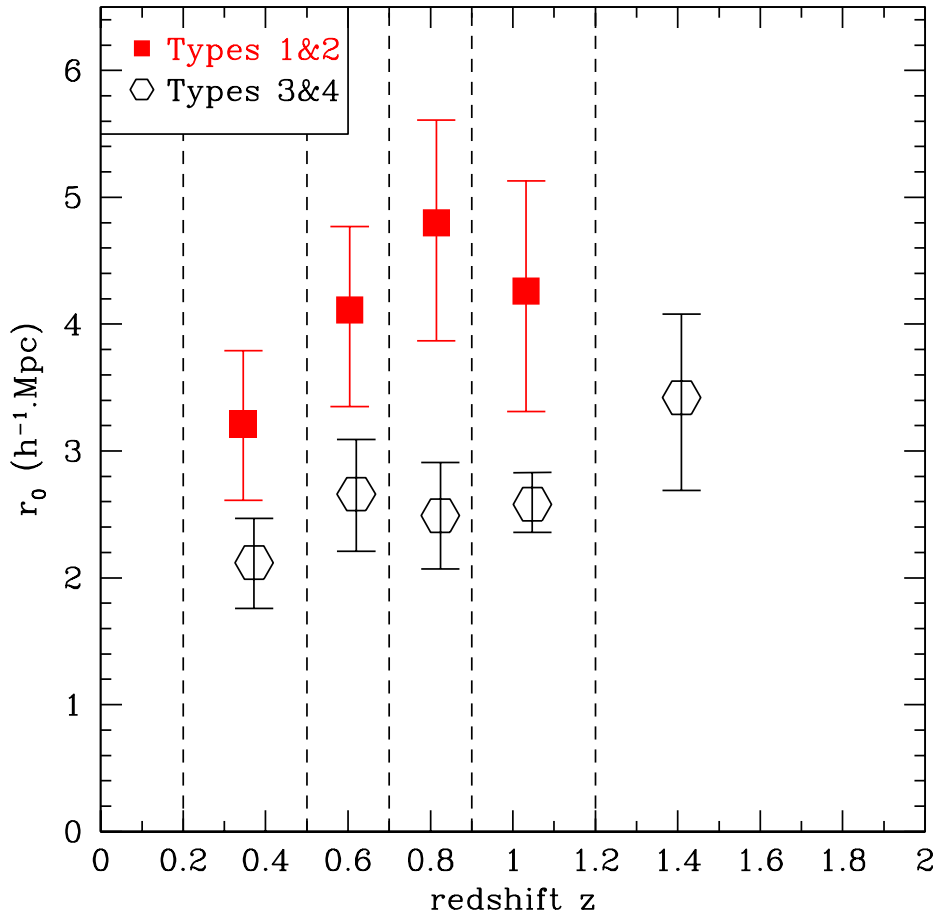


Fig. 8. Correlation length r_0 as a function of redshift per spectral type. Here the slope γ is a free parameter.

$z \simeq 1$ and with some indication for a diminishing trend beyond $z=1$. This result agrees with what has been found in a previous VVDS study (Marinoni et al. 2005), from the same data but using a different method. We also find a relative bias $b = 1.2 - 1.5$ (see table 3) between T1+T2 and T3, significantly lower than between T1+T2 and T4 and slightly increasing up to $z=1$.

5. Rest-frame color bimodality

It has become clear in recent years that the galaxy rest-frame colors show a bimodal distribution (Strateva et al. 2001), i.e. that in a (rest-frame) color-magnitude diagram galaxies tend to segregate between a “red sequence” (similar to, but less tight than that observed for cluster galaxies) and a “blue cloud”. This behaviour seems also to be present at high redshift (Bell et al. 2004; Giallongo et al. 2005). From the SDSS data, Strateva et al. (2001) find that the blue cloud contains mainly late (spiral) morphological types while the bulk of the red sequence consists of bulge-dominated, early-type galaxies (see also Weiner et al. 2005). This provides a natural recipe to split galaxies into two populations of red and blue objects and there have been already attempts to study the clustering of galaxies separately for the

two classes defined in this way at intermediate redshifts (*e.g.* Phleps et al. 2005).

It is therefore interesting to check how our analysis as a function of SED-defined types compares to a simpler subdivision following the bimodal distribution.

We have thus split our data into a blue and red sample, following the rest-frame color-magnitude relations suggested by Giallongo et al. (2005) and measured the projected function $w_p(r_p)$ and the best-fit power-law correlation function $\xi(r)$ in different redshift bins. The resulting evolution of the correlation length is shown in Fig. 10, with the measured values for r_0 and γ summarized in Table 5. The associated errors have been computed in the usual way, constructing blue and red mock catalogues from the GalICS simulations, keeping the same proportion of red (blue) objects as in the real VVDS data.

Similarly to what we found for early-type classes, red-sequence galaxies exhibit a larger clustering length ($r_0 \sim 4 h^{-1} \text{Mpc}$), with little dependence on redshift. Comparison to Fig. 8 indicates that the classification into rest-frame red and blue galaxies is substantially equivalent to our classification into spectral late-types, producing samples with similar clustering evolution properties. This is particularly true for $z < 0.7$. At higher redshift the sample of red-sequence galaxies shows

Table 2. Measurement of the clustering length for galaxy subsamples ranging from $z=0.2$ to $z=2$, letting the slope γ free or fixing the slope to the average value measured for a given type. The absolute $M_{B_{AB}}$ magnitude and number of objects are indicated.

type	number of galaxies	redshift range	effective redshift	$M_{B_{AB}} - 5\log(h)$		$r_0(z)$ ($h^{-1}\text{Mpc}$)	γ	$r_0(z)$ at fixed γ ($h^{-1}\text{Mpc}$)
1	164	[0.2-0.6]	0.427	-18.799	-19.020	$3.14^{+0.76}_{-0.94}$	$2.41^{+0.33}_{-0.37}$	$3.51^{+0.65}_{-0.71}$ ($\gamma = 2.15$)
	332	[0.6-1.0]	0.779	-19.755	-19.819	$4.35^{+0.68}_{-0.98}$	$2.01^{+0.46}_{-0.29}$	$3.93^{+0.55}_{-0.74}$
	237	[0.8-1.2]	0.964	-20.324	-20.347	$3.89^{+1.04}_{-1.44}$	$2.04^{+0.80}_{-0.45}$	$3.62^{+0.89}_{-1.26}$
2	299	[0.2-0.6]	0.409	-18.352	-18.419	$2.79^{+0.42}_{-0.44}$	$1.88^{+0.17}_{-0.14}$	$2.76^{+0.35}_{-0.32}$ ($\gamma = 1.90$)
	420	[0.6-1.0]	0.796	-19.525	-19.503	$4.92^{+0.72}_{-0.77}$	$1.82^{+0.27}_{-0.20}$	$4.89^{+0.58}_{-0.58}$
	358	[0.8-1.2]	0.972	-20.059	-20.115	$4.35^{+0.96}_{-1.14}$	$2.00^{+0.38}_{-0.26}$	$4.62^{+0.85}_{-0.96}$
3	437	[0.2-0.5]	0.346	-17.478	-17.332	$2.69^{+0.50}_{-0.54}$	$1.83^{+0.16}_{-0.11}$	$2.73^{+0.44}_{-0.42}$ ($\gamma = 1.78$)
	470	[0.5-0.7]	0.606	-18.665	-18.601	$3.21^{+0.60}_{-0.69}$	$1.91^{+0.24}_{-0.15}$	$3.33^{+0.54}_{-0.59}$
	473	[0.7-0.9]	0.810	-19.296	-19.212	$3.24^{+0.62}_{-0.65}$	$1.56^{+0.22}_{-0.15}$	$3.21^{+0.60}_{-0.60}$
	439	[0.9-1.2]	1.027	-19.843	-19.835	$2.88^{+0.44}_{-0.44}$	$1.83^{+0.36}_{-0.27}$	$2.82^{+0.49}_{-0.50}$
	191	[1.2-2.0]	1.394	-20.756	-20.751	$3.60^{+0.83}_{-1.38}$	$1.73^{+0.35}_{-0.46}$	$3.65^{+1.02}_{-3.04}$
4	507	[0.2-0.5]	0.366	-17.255	-17.170	$1.71^{+0.30}_{-0.31}$	$1.80^{+0.16}_{-0.12}$	$1.71^{+0.30}_{-0.30}$ ($\gamma = 1.86$)
	598	[0.5-0.7]	0.604	-18.277	-18.136	$2.31^{+0.40}_{-0.40}$	$1.76^{+0.14}_{-0.10}$	$2.28^{+0.42}_{-0.39}$
	534	[0.7-0.9]	0.807	-18.944	-18.829	$2.46^{+0.49}_{-0.57}$	$1.92^{+0.31}_{-0.19}$	$2.49^{+0.53}_{-0.58}$
	591	[0.9-1.2]	1.036	-19.553	-19.494	$2.85^{+0.36}_{-0.31}$	$1.94^{+0.18}_{-0.15}$	$2.91^{+0.40}_{-0.37}$
	289	[1.2-2.0]	1.394	-20.441	-20.397	$3.30^{+0.64}_{-0.87}$	$1.92^{+0.36}_{-0.32}$	$3.27^{+1.01}_{-2.35}$
1&2	307	[0.2-0.5]	0.346	-18.327	-18.512	$3.21^{+0.58}_{-0.60}$	$1.94^{+0.27}_{-0.29}$	$3.21^{+0.40}_{-0.40}$ ($\gamma = 1.89$)
	373	[0.5-0.7]	0.604	-18.996	-19.020	$4.11^{+0.66}_{-0.76}$	$1.94^{+0.19}_{-0.12}$	$4.14^{+0.53}_{-0.57}$
	390	[0.7-0.9]	0.814	-19.719	-19.635	$4.80^{+0.81}_{-0.93}$	$1.76^{+0.30}_{-0.18}$	$4.41^{+0.62}_{-0.69}$
	371	[0.9-1.2]	1.032	-20.333	-20.383	$4.26^{+0.87}_{-0.95}$	$1.92^{+0.31}_{-0.24}$	$4.32^{+0.79}_{-0.90}$
3&4	943	[0.2-0.5]	0.357	-17.358	-17.230	$2.12^{+0.35}_{-0.36}$	$1.78^{+0.14}_{-0.09}$	$2.12^{+0.32}_{-0.32}$ ($\gamma = 1.76$)
	1068	[0.5-0.7]	0.605	-18.448	-18.321	$2.66^{+0.43}_{-0.45}$	$1.74^{+0.15}_{-0.09}$	$2.66^{+0.41}_{-0.41}$
	1007	[0.7-0.9]	0.809	-19.109	-19.008	$2.49^{+0.42}_{-0.42}$	$1.67^{+0.17}_{-0.11}$	$2.46^{+0.48}_{-0.50}$
	1030	[0.9-1.2]	1.032	-19.677	-19.626	$2.58^{+0.25}_{-0.22}$	$1.86^{+0.11}_{-0.08}$	$2.58^{+0.28}_{-0.25}$
	480	[1.2-2.0]	1.394	-20.566	-20.510	$3.42^{+0.66}_{-0.73}$	$1.99^{+0.24}_{-0.20}$	$3.24^{+0.90}_{-1.23}$

Table 3. relative bias between early-type and late-type galaxies using spectral type classification, and “red” vs. “blue” galaxies separated using the color bi-modality

	redshift range	relative bias
T1&2/T3	[0.2-0.5]	$1.18^{+0.23}_{-0.22}$
	[0.5-0.7]	$1.27^{+0.26}_{-0.28}$
	[0.7-0.9]	$1.39^{+0.26}_{-0.27}$
	[0.9-1.2]	$1.45^{+0.34}_{-0.37}$
T1&2/T4	[0.2-0.5]	$1.76^{+0.27}_{-0.27}$
	[0.5-0.7]	$1.69^{+0.30}_{-0.30}$
	[0.7-0.9]	$1.85^{+0.43}_{-0.47}$
	[0.9-1.2]	$1.47^{+0.37}_{-0.40}$
“red”/“blue”	[0.2-0.5]	$1.69^{+0.43}_{-0.47}$
	[0.5-0.7]	$1.63^{+0.32}_{-0.34}$
	[0.7-0.9]	$1.45^{+0.29}_{-0.30}$
	[0.9-1.2]	$1.45^{+0.27}_{-0.27}$

a slightly weaker clustering than an early-type T1&2 selection. This is explained by noting that at $z > 0.7$ the red-sequence sample does include a significant fraction of late-type T3&4 galaxies (see Tables 2 and 5): 12% of the galaxies at $0.7 \leq z \leq 0.9$ are found to be late-type T3&4 objects with a mean absolute magnitude of $M_B \sim -19.3 + 5 \log(h)$ (see Table 4). The net effect of these galaxies is to lower the measured correlation signal.

Similarly, the bias between blue and red galaxies is measured to be varying from 1.7 to 1.45 from $z \sim 0.3$ to $z \sim 1$ as

presented in Table 3, and is therefore comparable to the bias measurement using the spectral type classification.

In summary, using the bi-modal split of the population we find results consistent with our analysis per spectral types. However, we note that the simple rest-frame color-magnitude cut is a less precise way of selecting truly early-type galaxies, at least for $z > 0.7$, since the full multi-color information is used to provide the spectral type classification.

Table 4. Number of galaxies and mean B rest-frame absolute magnitude ($h=1$) of the sample separated using the color bimodality. The number of “red” and “blue” galaxies classified as early (T1&2) and late (T3&4) spectral types are identified. We note that the populations of “red” and T1&2 as well as the populations of “blue” and T3&4 galaxies are mostly the same, except for a significant population cross-over which could explain the observed difference in clustering (see text).

redshift range	red		blue	
	T1&2	T3&4	T1&2	T3&4
0.2 - 0.5	221	49	86	895
	-18.181	-16.403	-18.701	-17.411
0.5 - 0.7	253	24	121	1044
	-18.892	-17.646	-19.215	-18.467
0.7 - 0.9	313	42	77	965
	-19.713	-18.468	-19.748	-19.138
0.9 - 1.2	283	13	88	1016
	-20.263	-19.314	-20.557	-19.679
1.2 - 2.0	41	0	94	480
	-20.926	0	-21.078	-20.567

Table 5. Measurements of the clustering length and slope for galaxy subsamples divided following the color bimodality

sample	number of galaxies	redshift range	effective redshift	$M_{BAB} - 5\log(h)$ mean	$r_0(z)$ ($h^{-1}\text{Mpc}$)	γ
red	270	0.2 - 0.5	0.345	-17.858	$3.84^{+0.87}_{-0.93}$	$1.88^{+0.23}_{-0.16}$
	277	0.5 - 0.7	0.610	-18.784	$4.35^{+0.65}_{-0.71}$	$1.99^{+0.22}_{-0.18}$
	355	0.7 - 0.9	0.809	-19.565	$3.78^{+0.70}_{-0.74}$	$1.87^{+0.28}_{-0.22}$
	296	0.9 - 1.2	1.022	-20.221	$3.72^{+0.63}_{-0.67}$	$1.83^{+0.32}_{-0.32}$
blue	980	0.2 - 0.5	0.357	-17.521	$2.07^{+0.34}_{-0.39}$	$1.67^{+0.14}_{-0.09}$
	1164	0.5 - 0.7	0.604	-18.544	$2.55^{+0.39}_{-0.41}$	$1.68^{+0.15}_{-0.09}$
	1042	0.7 - 0.9	0.811	-19.183	$2.18^{+0.32}_{-0.30}$	$1.40^{+0.16}_{-0.11}$
	1105	0.9 - 1.2	1.035	-19.749	$2.49^{+0.28}_{-0.22}$	$1.84^{+0.14}_{-0.10}$
	574	1.2 - 2.0	1.404	-20.651	$3.24^{+0.63}_{-0.70}$	$1.91^{+0.20}_{-0.18}$

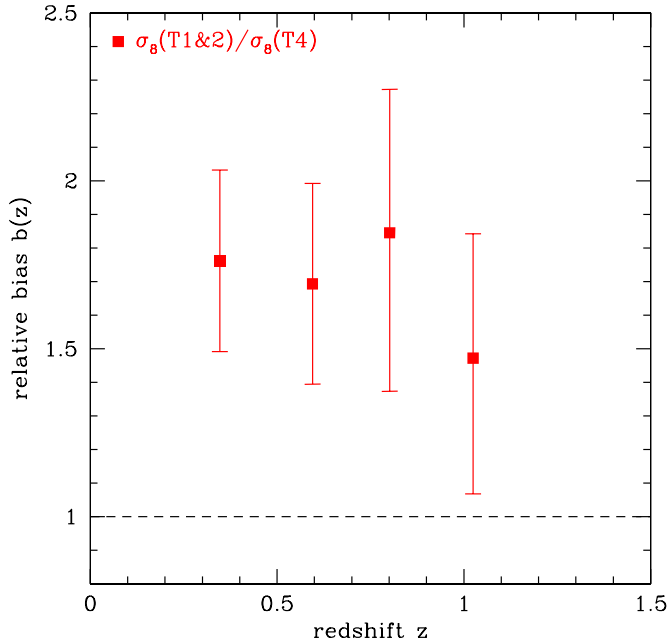


Fig. 9. relative bias between populations of galaxies for a $8\ h^{-1}\text{Mpc}$ scale: $b_A/b_B = \sigma_8^A/\sigma_8^B$.

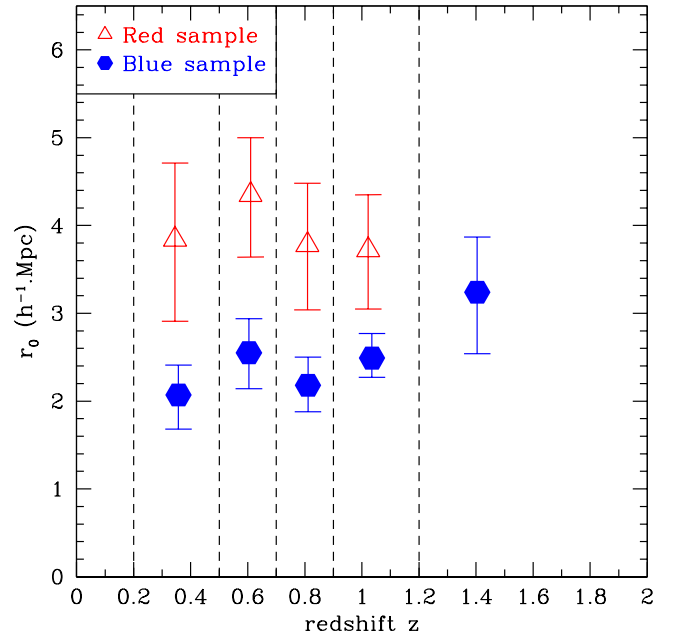


Fig. 10. Correlation length r_0 as a function of redshift for the red and blue samples. Here the slope γ is a free parameter.

6. Discussion and conclusions

The fundamental result of this investigation is to have clearly established that at least up to $z \approx 1.2$ early-type galaxies con-

tinue to be more strongly clustered than late-type galaxies. This

extends to higher redshifts earlier results obtained in the local Universe (Loveday Tresse & Maddox 1999; Norberg et al. 2002; Zehavi et al. 2002, 2005), at redshifts $z \simeq 0.6$ (Phleps et al. 2005), and quantitatively confirms early indications at $z \simeq 1$ (Phleps & Meisenheimer 2003).

Our specific results can be summarised as follows:

1. The clustering of sub- L^* early-type galaxies is stronger than the clustering of late-type galaxies up to $z \simeq 1.2$; a similar behaviour is obtained for “red-sequence” and “blue-cloud” galaxies selected using rest-frame colors.
2. Luminous, early-type galaxies are already strongly clustered at $z \sim 1$ with $4.5 \leq r_0 \leq 5 \text{ h}^{-1}\text{Mpc}$.
3. The clustering of late-type galaxies is low at all epochs with $r_0 \simeq 2.1 - 3.5 \text{ h}^{-1}\text{Mpc}$.
4. The overall clustering of bright late-type galaxies at $z \simeq 1.4$ is relatively strong with $r_0 = 3.4 \text{ h}^{-1}\text{Mpc}$.
5. The bias between the early-type and late type population is nearly constant, $b_{\text{early}}/b_{\text{late}} = 1.7$ over the redshift range $0.2 \leq z \leq 1.2$, with some indication of a decline beyond this redshift.

A physical interpretation of these results and comparison to local values, requires keeping in mind that we are working with a flux-limited sample, and thus that we are measuring galaxies of different intrinsic luminosity at different redshifts. At $z \sim 1$ the VVDS is observing early-type galaxies with $-22 \leq M_B(AB) \leq -19$ and an average $M_B(AB) = -20.3$ (see Figure 3). Early-type galaxies observed in the 2dFGRS with similar luminosities have an observed $r_0 = 6 - 7.5 \text{ h}^{-1}\text{Mpc}$, increasing with luminosity (Norberg et al. 2002). The clustering amplitude of $\xi(r)$ for early-type galaxies has therefore evolved by no more than 70% between redshift ~ 1 and present.

To understand the implications of this observation, we have compared our observed evolution of r_0 to the predicted evolution of the clustering length for the dark matter in a Λ cosmology (using the expected evolution of the variance as described in Mo & White (2002)). As shown in Figure 11, if one assumes that the early-type population is simply tracing passively the growth of clustering in the mass (as would be the case if no evolution in number nor in luminosity occurred) their clustering would increase by $\sigma_8(z=1)/\sigma_8(z=0) \simeq 0.6$, which is close to what we see in Figure 11. This result is in agreement with the observation of marginal evolution in the number density of early-type galaxies since $z \sim 1.5$ (Cimatti et al. 2004; Saracco et al. 2004), i.e. with the idea that the bulk of the spheroidal population was already in place at these early redshifts. If a significant fraction of these galaxies was assembled and added below $z \sim 1$, we would observe not only an increase in their number density, but also a weakening of their clustering signal with respect to the underlying mass: it is natural to think that, forming typically in lower density peaks than their high-redshift companions they would in fact be less massive and less clustered.

The lower clustering level observed for early-type galaxies at $z < 0.7$, is best interpreted as a luminosity effect, given the faint typical luminosity of galaxies in the first bins (*e.g.* $\langle M_B(AB) \rangle = -18.3$ in the bin at $z = [0.2, 0.5]$).

The trend observed for blue star-forming galaxies of similar luminosity is also very interesting. Below redshift $z \simeq 1$, their clustering is observed to be weak and continuously de-

creasing from high redshift to the present. As shown in the two right-top panels of Figure 11, however, this is consistent with simple gravitational growth: the clustering amplitude measured at $z \sim 1$ for these galaxies is coherent with the evolution of the same structure traced by the same galaxies at $z \sim 0$. Only when we examine the point at $z = 1.4$, we see a discrepancy with local values, indicating that star formation activity at $z \gg 1$ is found in progressively more strongly clustered galaxies than at present, i.e. plausibly in more massive galaxies in higher density peaks. In our sample, galaxies at $z = [1.2, 2]$ have a mean $M_B(AB) = -20.5$ and a clustering length $r_0 = 3.5 \text{ h}^{-1}\text{Mpc}$. Taking into account 1.5 magnitudes of B-band brightening (Zucca et al. 2005) these galaxies would have today $M_B(AB) = -19$. Late-type galaxies with similar luminosities in the Stromlo-APM, 2dFGRS or SDSS have $r_0 \simeq 2.9 - 4.2 \text{ h}^{-1}\text{Mpc}$ (Loveday Tresse & Maddox 1999; Norberg et al. 2002; Zehavi et al. 2002). The clustering amplitude of late-type star-forming galaxies with comparable luminosity, therefore remains roughly constant since $z \sim 1.5$ (Figure 11). A straightforward explanation to produce the observed result is that as redshifts get smaller, star formation is shifting to smaller-mass, less clustered galaxies.

Comparing our values with other clustering measurements for type- or color- selected galaxies at redshifts beyond 1.5, we find for example that star-forming LBGs at $z \sim 3$ (Adelberger et al. 2004) are even more clustered ($r_0 \sim 4.5 \text{ h}^{-1}\text{Mpc}$), than our star-forming galaxies at $z \sim 1.5$. This is consistent with the VVDS results at lower redshifts, indicating that star formation is moving to higher and higher density environments, when moving back in time. Similarly, red selected samples of EROs at $z \sim 1.5$ ERO’s or red K selected galaxies with $2 < z_{\text{phot}} < 4$ are significantly more clustered than the early type galaxies at $z \simeq 1$ in the VVDS with r_0 up to $8 \text{ h}^{-1}\text{Mpc}$ (Daddi et al. 2002, 2003). These galaxies seem to be populating even more extreme density peaks than those we have been studying here at $z \sim 1$.

Acknowledgements. This research program has been developed within the framework of the VVDS consortium.

This work has been partially supported by the CNRS-INSU and its Programme National de Cosmologie (France), and by Italian Ministry (MIUR) grants COFIN2000 (MM02037133) and COFIN2003 (num.2003020150).

The VLT-VIMOS observations have been carried out on guaranteed time (GTO) allocated by the European Southern Observatory (ESO) to the VIRMOS consortium, under a contractual agreement between the Centre National de la Recherche Scientifique of France, heading a consortium of French and Italian institutes, and ESO, to design, manufacture and test the VIMOS instrument.

References

- Adelberger, K.L., Steidel, C.C., Pettini, M., Shapley, A.E., Reddy, N.A. & Erb, D.K. 2005, *ApJ*, 619, 697
- Arnouts, S., Cristiani, S., Moscardini, L. et al. 1999, *MNRAS*, 310, 540
- Bell, E. F., Wolf, C., Meisenheimer, K. et al., 2004 *ApJ*, 608, 752

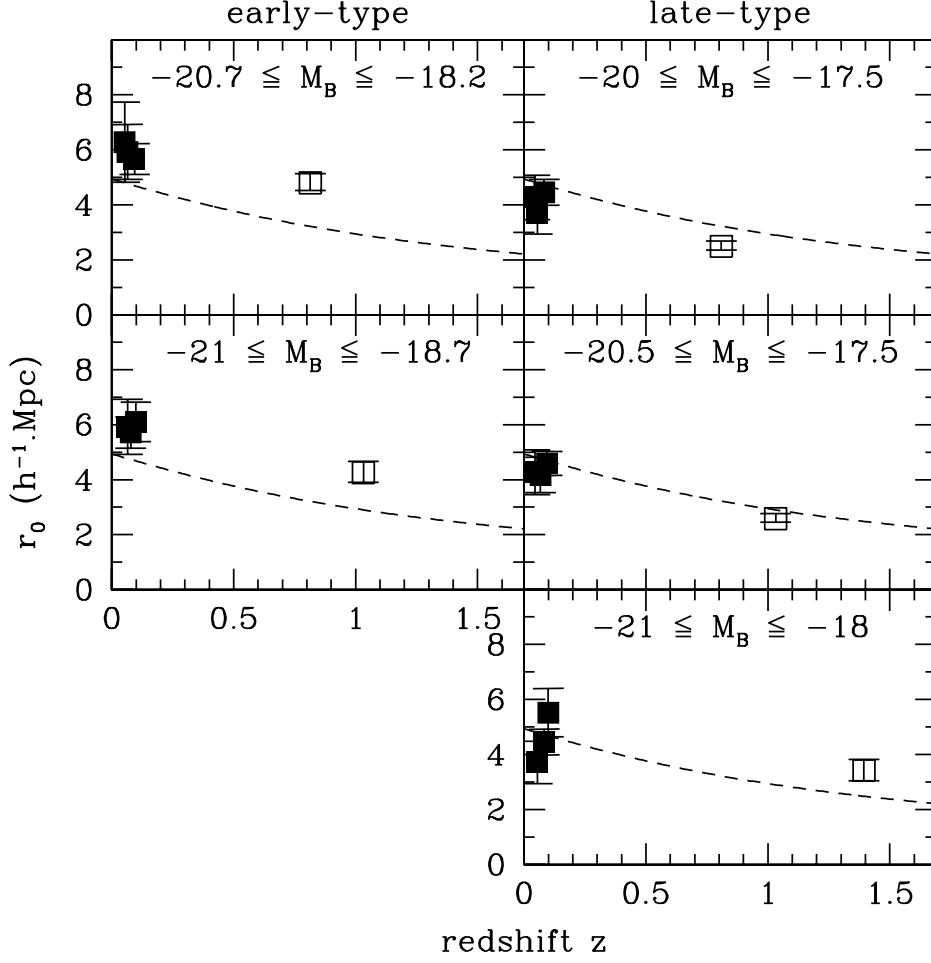


Fig. 11. Clustering length evolution for early (left panel) and late (right panel) spectral type galaxies. Measurements from the VVDS at high redshifts (open symbols) are compared to the local measurements obtained for galaxies in the same absolute luminosity ranges in the 2dFGRS (filled symbols) surveys (Norberg et al. 2002), with an evolution correction of 0.5 magnitudes for early types and 1 magnitude for late types at $z \sim 1$ (Zucca et al. 2005). The solid curves show the predicted gravitational growth of the dark matter, with amplitude each time normalised to the VVDS measurements.

- Benoist, C., Maurogordato, S., da Costa, L. N., Cappi, A. & Schaeffer, R. 1996, *ApJ*, 472, 452
- Blaizot, J., Wadadekar, Y., Guiderdoni, B. et al. 2005, *MNRAS*, 360, 159
- Bruzual, G. & Charlot, S. 1993, *ApJ*, 405, 538
- Carlberg, R. G., Cowie, L. L., Songaila, A. & Hu, E. M. 1997, *ApJ*, 484, 538
- Cimatti, A., Daddi, E., Renzini, A. et al. 2004, *Nature*, 430, 184
- Coil, A.L., Davis, M., Madgwick, D.S. et al. 2004, *ApJ*, 609, 525
- Coleman, G.D., Wu, C. & Weedman, D.W. 1980, *ApJS*, 43, 393
- Daddi, E., Cimatti, A., Broadhurst, T. et al. 2002, *A&A*, 384, L1
- Daddi, E., Rottgering, H., Labbe, I. et al. 2003, *ApJ*, 588, 50
- Davis, M., Peebles, J.P.E. 1983, *ApJ*, 267, 465
- Dressler, A. 1980, *ApJ*, 236, 351
- Firth, A. E., Somerville, R.S., McMahon, R.G. et al. 2002, *MNRAS*, 332, 617
- Gaztañaga, E., & Juszkiewicz, R. 2001, *ApJ*, 558, L1
- Giallongo, E., Salimbeni, S., Menci, N. et al. 2005, *ApJ*, 622, 116
- Guzzo, L., Iovino, A., Chincarini, G., Giovanelli, R., & Haynes, M. P. 1991, *ApJ*, 382, L5
- Guzzo, L., Strauss, M.A., Fisher K.B., Giovanelli, R., Haynes, M.P. 1997, *ApJ*, 489, 37
- Ilbert, O., Tresse, L., Zucca, E., Bardelli, S. and the VVDS team 2005, *A&A*, 439, 863
- Ilbert, O., Bolzonella, M., and the VVDS team 2005, in preparation
- Iovino, A., Melnick, J., Shaver, P. 1988, *ApJ*, 330, L17
- Iovino, A., Giovanelli, R., Haynes, M., Chincarini, G. & Guzzo, L. 1993, *MNRAS*, 265, 21
- Jenkins, A., Frenk, C.S., Pearce, F.R. et al. 1998, *ApJ*, 499, 20
- Kaiser, N. 1984, *ApJ*, 284, L9
- Kauffmann, G., Colberg, J.M., Diaferio, A. & White, S.D.M. 1999, *MNRAS*, 303, 188

- Kauffmann, G., Colberg, J.M., Diaferio, A. & White, S.D.M. 1999, MNRAS, 307, 529
- Lahav, O., et al. 2002, MNRAS, 333, 961
- Landy, S.D. and Szalay, A.S. 1993, ApJ, 412, 64
- Le Fèvre, O., Hudon, D., Lilly, S. J. et al. 1996, ApJ, 461, 534
- Le Fèvre, O., Mellier, Y., McCracken, H.J. et al. 2004a, A&A, 417, 839
- Le Fèvre, O., Vettolani, G., Garilli, B. et al. 2005, A&A, 2005, 439, 845
- Le Fèvre, O., and the VVDS team 2004, A&A, 428, 1043
- Le Fèvre, O., Guzzo, L., Meneux, B. et al. 2005, A&A, 2005, 439, 877
- Loveday, J., Tresse, L. & Maddox, S. 1999, MNRAS, 310, 281
- Madgwick, D.S., Hawkins, Ed, Lahav, O. et al. 2003, MNRAS, 344, 847
- Marinoni, C., Le Fèvre, O., Meneux, B. and the VVDS team 2005, A&A, 442, 801
- McCracken, H.J., Radovich, M., Bertin, E. et al. 2003, A&A, 410, 17
- Mo, H. J. & White, S. D. M. 1996, MNRAS, 282, 347
- Mo, H. J. & White, S. D. M. 2002, MNRAS, 336, 112
- Norberg, P., Baugh, C.M., Hawkins, Ed et al. 2002, MNRAS, 332, 827
- Peebles, P. J. E. 1980, The Large Scale Structure of the Universe (Princeton University Press, Princeton)
- Phleps, S. & Meisenheimer, K. 2003, A&A, 407, 855
- Phleps, S., Peacock, J. A., Meisenheimer, K. & Wolf, C. 2005, astro-ph/0506320
- Pollo, A., Meneux, B., Guzzo, L., Le Fèvre, O., and the VVDS team 2005 A&A, 439, 887
- Postman, M., Franx, M., Cross, N. J. G. et al. 2005, ApJ, 623, 721
- Radovich, M., Arnaboldi, M., Ripepi, V., et al. 2004, A&A, 417, 51
- Saracco, P., Longhetti, M.; Giallongo, E. et al. 2004, A&A, 420, 125
- Sheth, R. K. & Tormen, G. 1999, MNRAS, 308, 119
- Shepherd, C. W., Carlberg, R. G., Yee, H. K. C. et al. 2001, ApJ, 560, 72
- Smith, G.P., Treu, T., Ellis, R.S., Moran, S.M., Dressler, A. 2005, ApJ, 620, 78
- Springel, V., White, S.D.M., Jenkins, A. et al. 2005, Nature, 435, 629
- Strateva, I., Ivezić, Z., Knappet, G.R. et al. 2001, AJ, 122, 1861
- van Dokkum, P. G., Franx, M., Fabricant, D., Illingworth, G. D., & Kelson, D. D. 2000, ApJ, 541, 95
- Weiner, B.J., Phillips, A.C., Faber, S.M. et al. 2005, ApJ, 620, 595
- Willmer, C.N.A., et al. 2005, A&A, astro-ph/0506041
- Zehavi, I., Blanton, M.R., Frieman, J.A. et al. 2002, ApJ, 571, 172
- Zehavi, I., Weinberg, D.H., Zheng, Z. et al. 2004, ApJ, 608, 16
- Zehavi, I., Zheng, Z., Weinberg, D.H. et al. 2005, ApJ, 630, 1
- Zucca, E., Ilbert, O., Bardelli, S. and the VVDS team 2005, A&A, astro-ph/0506393

5.3 Évolution de la fonction de corrélation pour un échantillon de galaxies sélectionnées dans l'ultraviolet

Il est important de mesurer l'état d'agrégation de galaxies sélectionnées en ultraviolet. En effet, si ces dernières sont visibles à ces longueurs d'ondes, cela signifie qu'elles sont en train de former beaucoup d'étoiles. Étudier la fonction de corrélation de ces galaxies nous permet donc de déterminer et de quantifier où a lieu la formation stellaire : dans une région dense si l'agrégation est forte, dans le champ si la longueur de corrélation est faible. Étudier l'évolution de l'agrégation de ces populations de galaxies permet de savoir si la formation stellaire s'est déplacée au cours du temps entre des régions de différentes densités ou non.

5.3.1 Galex : observation en NUV et FUV

Le satellite GALEX ([Martin et al., 2005](#)) a été lancé par la NASA le 28 avril 2003. Il s'agit d'une mission spatiale destinée à retracer l'histoire et la formation stellaire globale sur 80% de l'âge de l'Univers. Au cours de cette période les galaxies ont évolué considérablement, et la majeure partie des éléments, étoiles et disques galactiques ont ainsi été formés. L'observation en ultraviolet permet d'observer ces galaxies qui sont justement en train de former des étoiles.

Le champ VVDS-02h a été imagé en ultraviolet proche (NUV, $\sim 2310\text{\AA}$) et lointain (FUV, $\sim 1530\text{\AA}$) dans le cadre du relevé profond *GALEX Deep Imaging survey*. La résolution spatiale de GALEX étant de $\sim 5''$ (contre $\sim 0.2''$ pour les images optiques), 52% des sources $\text{NUV} \leq 24.5$ ont ainsi au moins deux contreparties optiques ([Arnouts et al., 2005](#)).

On associe à chaque source UV sa (plus proche) contrepartie optique (lorsqu'il y en a plusieurs). On définit ainsi deux échantillons spectroscopiques sélectionnés en NUV et FUV. Les comptages des catalogues photométriques et spectroscopiques en NUV et FUV sont représentés figure 5.10. Les échantillons sont complets à 80% jusqu'aux magnitudes 24.5 ([Arnouts et al., 2005](#); [Xu et al., 2005](#)), ainsi que le montre sous une autre forme la figure 5.11. De par la limite $I_{AB} = 24$, on estime qu'une fraction de $\sim 4\%$ d'objets manquent aux catalogues spectroscopiques. Il s'agit d'objets bleus et faibles en ultraviolet. [Arnouts et al. \(2005\)](#) estiment encore que $\sim 95\%$ des objets saturés en I_{AB} (≤ 17.5) mais détectés en NUV sont situés à bas *redshifts* (≤ 0.2).

La figure 5.12 représente les distributions en *redshifts* des catalogues de galaxies plus brillantes que $\text{NUV}=24$ et 24.5 , $\text{FUV}=24$ et 24.5 . On observe ainsi des structures en FUV jusqu'à $z \sim 0.5$ et en NUV jusqu'à $z \sim 1.2-1.5$. On ne détecte plus d'objets dès lors que la discontinuité de Lyman à 912\AA sort de la plage des filtres en ultraviolet puisque les galaxies n'émettent presque plus aucun flux en-deçà de cette limite.

Comme attendu, une très large majorité des 4 échantillons VVDS-GALEX correspondent à des galaxies de type spectral tardif 3 ou 4 (voir la table récapitulative 6.4 du chapitre 6). Cette proportion passe de $\sim 80\%$ pour les 4 échantillons à bas *redshift*, jusqu'à 90% en FUV à $z \sim 0.4$ et 99% en NUV à $z \sim 0.9$.

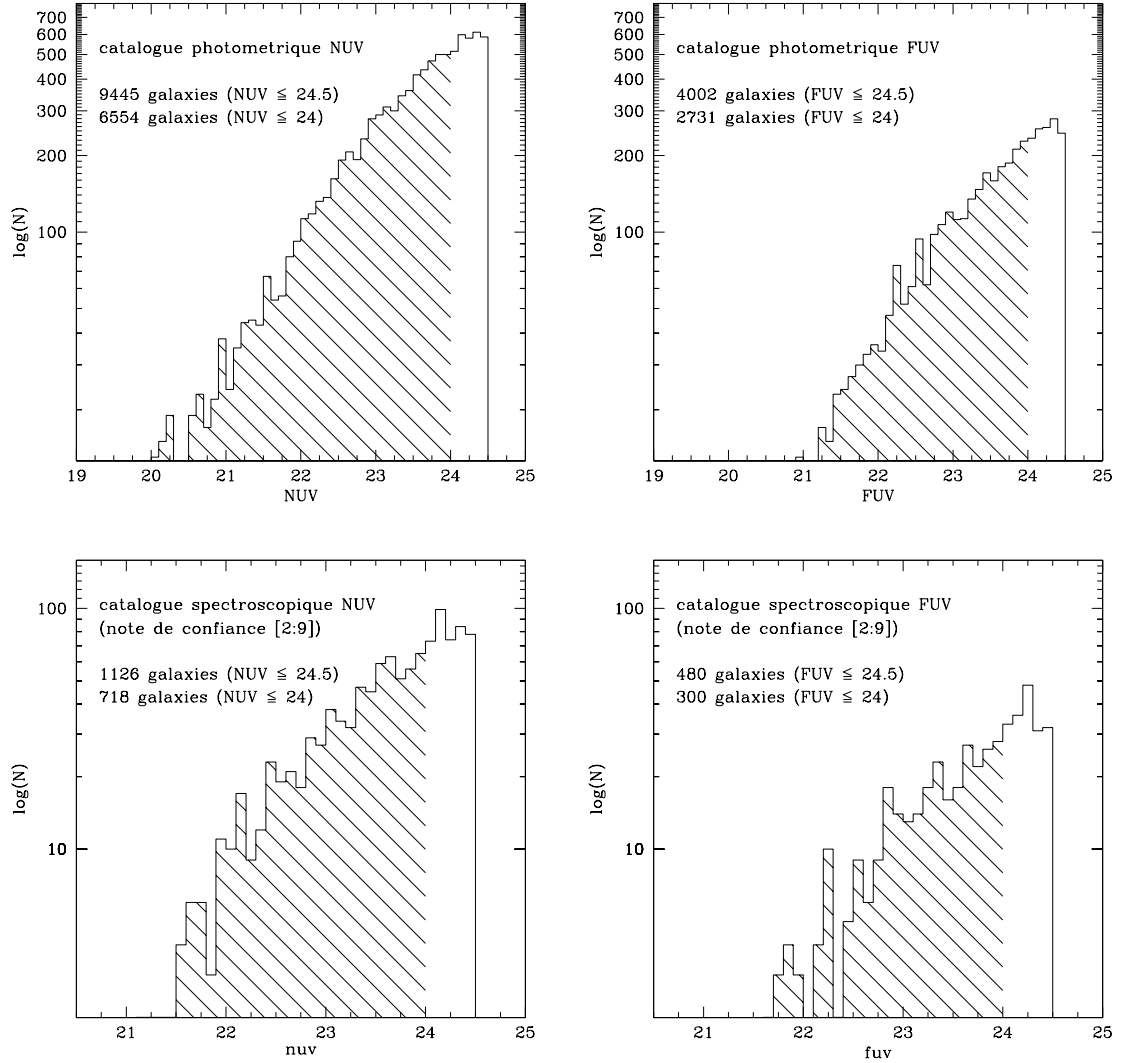


FIG. 5.10 – Comptages en NUV (à gauche) et FUV (à droite)

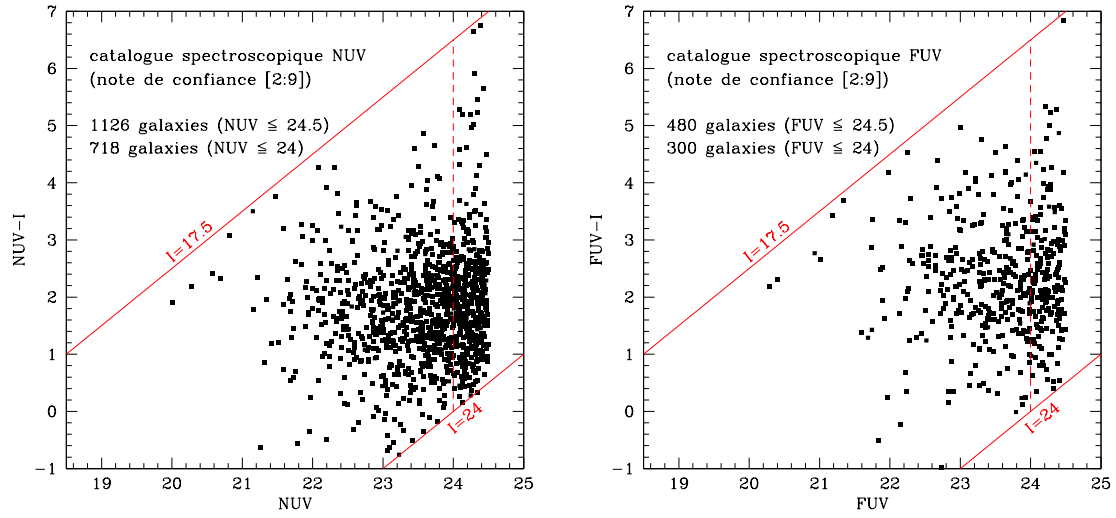
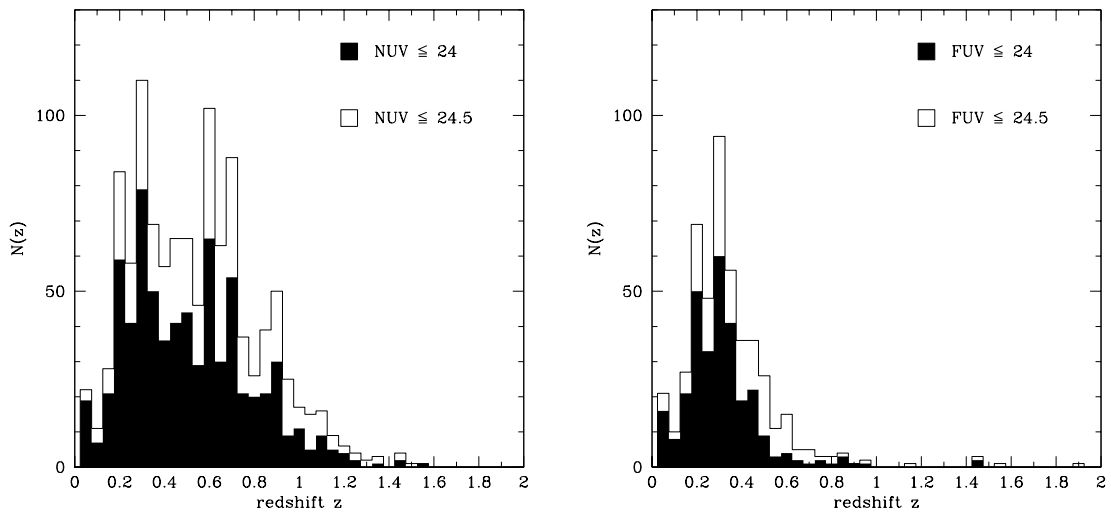


FIG. 5.11 – Diagramme couleur-magnitude en NUV (à gauche) et FUV (à droite)

FIG. 5.12 – Distributions en *redshift* des échantillons GALEX-VVDS sélectionnés en NUV (à gauche) et en FUV (à droite)

5.3.2 Fonction de corrélation en NUV

Je présente ici la mesure des paramètres de la fonction de corrélation pour l'échantillon sélectionné en NUV dans trois intervalles en *redshift*. La figure 5.13 montre la meilleure représentation de la fonction $w_p(r_p)$ obtenue entre $z=0.2$ et $z=1.2$, avec les contours d'erreurs associés dans le plan (r_0, γ) . Les erreurs indiquées ici sont celles obtenues à partir d'une simple méthode de rééchantillonnage *bootstrap*. Les mesures sont résumées dans le tableau 5.2.

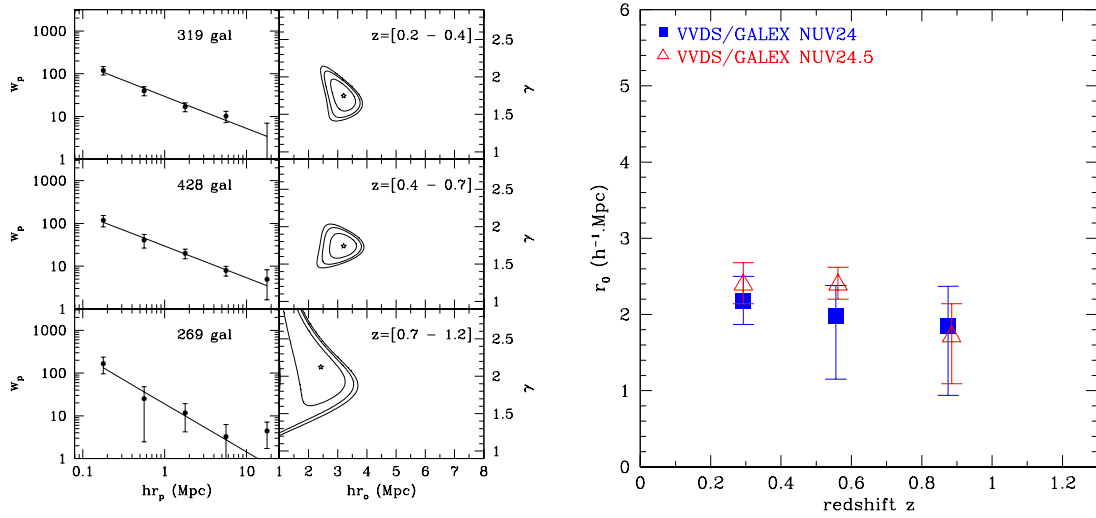


FIG. 5.13 – Fonction de corrélation $w_p(r_p)$ mesurée à partir de l'échantillon GALEX-VVDS en $\text{NUV} \leq 24.5$ (à gauche). On a représenté à droite l'évolution de la longueur de corrélation $r_0(z)$ pour $\text{NUV} \leq 24$ (carrés pleins) et $\text{NUV} \leq 24.5$ (triangles ouverts).

Qu'elle soit mesurée à $\text{NUV} \leq 24$ ou $\text{NUV} \leq 24.5$, la longueur de corrélation r_0 évolue peu entre $z=0.2$ et 1.2 alors même que les populations sondées sont intrinsèquement beaucoup plus brillantes (2 magnitudes d'écart en bande B au repos). De plus sa valeur reste faible et même inférieure ($r_0 \sim 2 \text{ h}^{-1} \text{Mpc}$) aux mesures réalisées pour les populations de type tardif T3&4 (section 5.2) qui dominent l'échantillon. Par contre, les pentes γ mesurées en NUV sont plus importantes. En moyenne, les contreparties optiques des objets sélectionnés en NUV sont une magnitude plus brillantes que l'ensemble de la population de galaxies de type T3&4 du VVDS. Et comme attendu, ces contreparties représentent un sous-échantillon en moyenne nettement plus bleu que la population globale (figure 5.14). Tout ceci indique que l'on observe des galaxies de champ présentant un fort taux de formation stellaire.

Échantillon	Nombre de galaxies	Intervalle en <i>redshift</i>	<i>redshift</i> moyen	r_0 ($h^{-1} \cdot \text{Mpc}$)	γ
$\text{NUV} \leq 24$	223	0.2 – 0.4	0.293	$2.18^{+0.32}_{-0.31}$	$2.08^{+0.29}_{-0.24}$
	263	0.4 – 0.7	0.556	$1.98^{+0.40}_{-0.83}$	$1.55^{+0.28}_{-0.31}$
	154	0.7 – 1.2	0.875	$1.85^{+0.52}_{-0.91}$	$2.43^{+1.73}_{-0.63}$
$\text{NUV} \leq 24.5$	320	0.2 – 0.4	0.293	$2.39^{+0.29}_{-0.25}$	$2.10^{+0.22}_{-0.19}$
	428	0.4 – 0.7	0.562	$2.39^{+0.23}_{-0.19}$	$2.09^{+0.16}_{-0.15}$
	269	0.7 – 1.2	0.885	$1.71^{+0.43}_{-0.62}$	$2.64^{+1.00}_{-0.45}$

TAB. 5.2 – Longueur de corrélation r_0 et pente γ de $\xi(r)$ mesurés à partir des échantillons GALEX-VVDS sélectionnés en NUV (Erreurs obtenues par rééchantillonnage “bootstrap”)

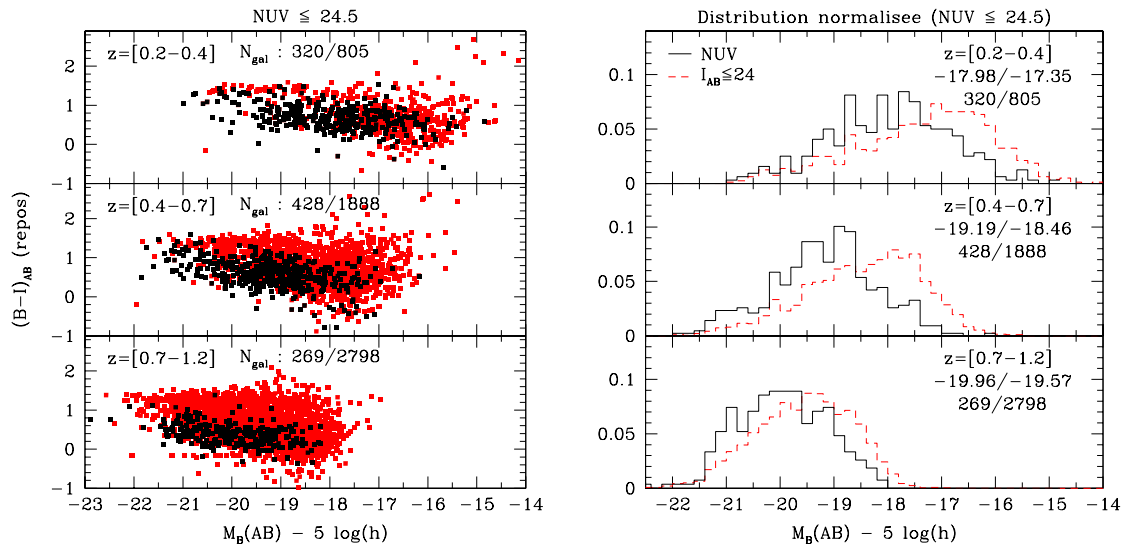


FIG. 5.14 – Couleurs en NUV

5.3.3 Fonction de corrélation en FUV

Je présente ici la mesure des paramètres de la fonction de corrélation pour l'échantillon sélectionné en FUV dans deux intervalles en *redshift*. La figure 5.15 montre la meilleure représentation de la fonction $w_p(r_p)$ obtenue entre $z=0.2$ et $z=1.2$, avec les contours d'erreurs associés dans le plan (r_0, γ) . Les erreurs indiquées ici sont celles obtenues à partir d'une simple méthode de rééchantillonnage “bootstrap”. Les mesures sont résumées dans le tableau 5.3.

On observe une très forte évolution de la longueur de corrélation r_0 mesurée en FUV. Il faut tout de même noter que l'échantillon à $z \sim 0.2$ est en moyenne une magnitude plus

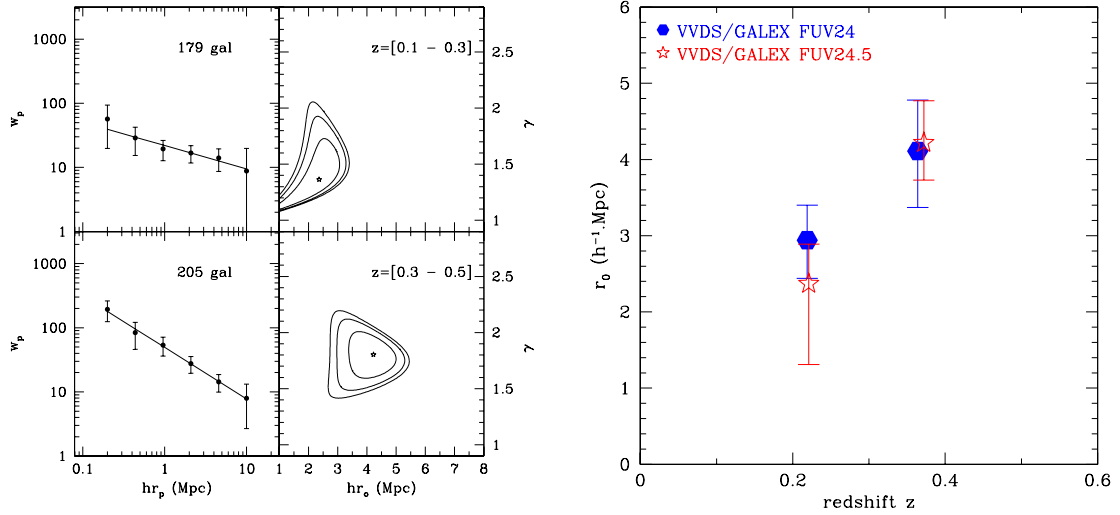


FIG. 5.15 – Fonction de corrélation $w_p(r_p)$ mesurée à partir de l'échantillon GALEX-VVDS en $FUV \leq 24.5$ (à gauche). On a représenté à droite l'évolution de la longueur de corrélation $r_0(z)$ pour $FUV \leq 24$ (hexagones pleins) et $FUV \leq 24.5$ (étoiles).

faible (en bande B au repos) que celui à $z \sim 0.4$. La encore, les contreparties optiques des objets détectés en FUV sont plus brillantes et plus bleues que la population globale des galaxies (figure 5.16). Ces objets, détectés en FUV par le flux émis par leur étoiles jeunes et chaudes de type OB, se trouvent dans des régions de pics de densité. La longueur de corrélation mesurée est en effet de l'ordre de $3-4 \text{ h}^{-1} \text{ Mpc}$. La formation stellaire est donc fortement influencée par l'environnement.

Échantillon	Nombre de galaxies	Intervalle en <i>redshift</i>	<i>redshift</i> moyen	r_0 ($\text{h}^{-1} \text{ Mpc}$)	γ
$FUV \leq 24$	128	0.1 – 0.3	0.219	$2.94^{+0.46}_{-0.50}$	$1.55^{+0.20}_{-0.17}$
	124	0.3 – 0.5	0.364	$4.11^{+0.67}_{-0.74}$	$1.85^{+0.32}_{-0.24}$
$FUV \leq 24.5$	179	0.1 – 0.3	0.221	$2.37^{+0.52}_{-1.06}$	$1.36^{+0.24}_{-0.20}$
	205	0.3 – 0.5	0.372	$4.22^{+0.55}_{-0.49}$	$1.80^{+0.14}_{-0.12}$

TAB. 5.3 – Longueur de corrélation r_0 et pente γ de $\xi(r)$ mesurés à partir des échantillons GALEX-VVDS sélectionnés en NUV (Erreurs obtenues par rééchantillonnage “bootstrap”)

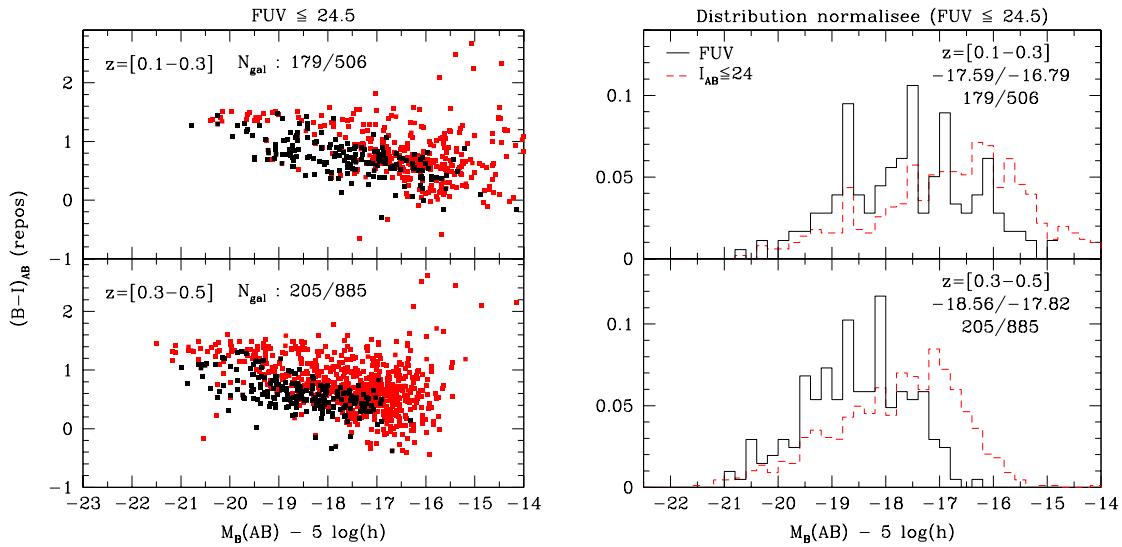


FIG. 5.16 – Couleurs en FUV

Ces résultats tout à fait nouveaux pour un échantillon sélectionné à ces longueurs d'onde sera soumis pour publication dans la revue européenne *Astronomy & Astrophysics*.

5.4 Évolution du biais non-linéaire jusqu'au *redshift* $z=1.5$

5.4.1 Résultats

Il est possible dans le cas d'un sondage profond comme celui du VVDS de calculer la fonction de probabilité de distribution des fluctuations (PDF) du champ de densité des galaxies. Le calcul de cette PDF à différents *redshifts* permet d'éclaircir les idées générales sur le développement des structures par l'effondrement gravitationnel des fluctuations de densité. [Marinoni et al. \(2005\)](#) ont ainsi calculé pour la première fois l'évolution de la PDF des galaxies à *redshift* $0.7 \leq z \leq 1.5$ pour un échantillon de galaxies de la région à 4 passages du VVDS-02h, limité en volume à $M_B < -20 + 5 \log(h)$. J'ai contribué à ce travail en collaboration avec Christian Marinoni, en particulier au niveau du calcul d'erreur en utilisant les simulations *GalICS*, comme décrit dans l'article.

La méthode de reconstruction du champ de densité des galaxies a été testée à partir de catalogues fictifs construits à partir des simulations *GalICS* (section 4.2) et *Hubble-Volume* (section 4.1.1). Une fois appliquée aux données du VVDS, les résultats obtenus sont les suivants :

- Les fluctuations du champ de densité des galaxies, σ_8 , restent relativement constante entre $z=1.5$ et $z=0.7$. Sa valeur moyenne, dans le cas de l'échantillon limité en volume est de 0.94 ± 0.07 .
- La probabilité de détecter une région sous-dense à $z \sim 0.7$ est plus forte qu'à $z \sim 1.5$.
- La comparaison des mesures faites de σ_8 avec des prédictions basées sur des modèles linéaires et non-linéaires indique de façon indubitable que le biais des galaxies est une fonction croissante du *redshift*.

- Le biais est une fonction non-linéaire entre $z=0.7$ et 1.5 et les effets de sa composante non linéaire sont de l'ordre de $\sim 10\%$ sur des échelles supérieures à $5 h^{-1} \text{Mpc}$. Ce résultat confirme ainsi les prédictions théoriques faites par les modèles hiérarchiques *Cold Dark Matter* de formation des galaxies (Somerville *et al.*, 2001, par exemple).
- La distribution des galaxies brillantes est plus fortement biaisée que celle des galaxies faibles.
- Le biais de la population de galaxies rouges ($(B-I) > 1.5$) est plus prononcé que celui de la population bleue ($(B-I) < 1.0$). On observe peu d'évolution du biais relatif entre les deux populations entre $z=0.7$ et $z=1.5$. Sa valeur moyenne est de 1.4 ± 0.1 .
- Des trois modèles de biais décrits dans la section 1.6, celui où le nombre de galaxies se conserve au cours du temps (Fry, 1996) et celui prenant en compte la fusion des galaxies (Mo et White, 1996) décrivent très mal les données du VVDS contrairement au troisième, qui fait l'hypothèse que les galaxies plus brillantes qu'une certaine luminosité L tracent bien la distribution des halos de masse M (Blanton *et al.*, 2000).

Tous ces résultats sont détaillés dans l'article de Marinoni *et al.* (2005), soumis pour publication dans la revue *Astronomy & Astrophysics* et présenté page section 5.4.2.

5.4.2 Article IV : *The VIMOS VLT Deep Survey - evolution of the non-linear galaxy bias up to $z=1.5$*

Cet article a été publié dans la revue *Astronomy & Astrophysics* par Marinoni, C., Le Fèvre, O., Meneux, B., *et al.* (2005, A&A, 442). Il présente la mesure et l'évolution du biais entre la distribution spatiale des galaxies et celle de la matière noire dans sondage VVDS jusqu'à $z \sim 2$.

The VIMOS VLT Deep Survey^{★,★★}

Evolution of the non-linear galaxy bias up to $z = 1.5$

C. Marinoni^{1,2}, O. Le Fèvre², B. Meneux², A. Iovino¹, A. Pollo³, O. Ilbert^{2,5}, G. Zamorani⁵, L. Guzzo³, A. Mazure², R. Scaramella⁴, A. Cappi⁵, H. J. McCracken⁶, D. Bottini⁷, B. Garilli⁷, V. Le Brun², D. Maccagni⁷, J. P. Picat⁸, M. Scodeggio⁷, L. Tresse², G. Vettolani⁹, A. Zanichelli⁹, C. Adami², S. Arnouts², S. Bardelli⁵, J. Blaizot², M. Bolzonella¹⁰, S. Charlot^{6,11}, P. Ciliegi⁹, T. Contini⁸, S. Foucaud⁷, P. Franzetti⁷, I. Gavignaud⁸, B. Marano¹⁰, G. Mathez⁸, R. Merighi⁵, S. Paltani², R. Pellò⁸, L. Pozzetti⁵, M. Radovich¹², E. Zucca⁵, M. Bondi⁹, A. Bongiorno¹⁰, G. Busarello¹², S. Colombi⁶, O. Cucciati^{1,13}, F. Lamareille⁸, Y. Mellier⁶, P. Merluzzi¹², V. Ripepi¹², and D. Rizzo⁸

¹ INAF - Osservatorio Astronomico di Brera, via Brera 28, 20121 Milano, Italia

² Laboratoire d'Astrophysique de Marseille, UMR 6110 CNRS-Université de Provence, Traverse du Siphon-Les trois Lucs, 13012 Marseille, France
e-mail: christian.marinoni@cpt.univ-mrs.fr

³ INAF - Osservatorio Astronomico di Brera, via Bianchi 46, 23807 Merate, Italia

⁴ INAF - Osservatorio Astronomico di Roma, via Osservatorio 2, 00040 Monteporzio Catone (Roma), Italia

⁵ INAF - Osservatorio Astronomico di Bologna, via Ranzani 1, 40127 Bologna, Italia

⁶ Institut d'Astrophysique de Paris, UMR 7095, 98 bis Bd Arago, 75014 Paris, France

⁷ INAF - IASF, Via Bassini 15, 20133 Milano, Italia

⁸ Laboratoire d'Astrophysique - Observatoire Midi-Pyrénées, Toulouse, France

⁹ INAF - Istituto di Radio-Astronomia, Via Gobetti 101, 40129 Bologna, Italia

¹⁰ Università di Bologna, Dipartimento di Astronomia, via Ranzani 1, 40127 Bologna, Italia

¹¹ Max Planck Institut für Astrophysik, 85741 Garching, Germany

¹² INAF - Osservatorio Astronomico di Capodimonte, via Moiariello 16, 80131 Napoli, Italia

¹³ Università di Milano-Bicocca, Dipartimento di Fisica, Piazza della scienza 3, 20126 Milano, Italia

Received 2 March 2005 / Accepted 17 June 2005

ABSTRACT

We present the first measurements of the Probability Distribution Function (PDF) of galaxy fluctuations in the four-passes, first-epoch VIMOS-VLT Deep Survey (VVDS) cone, covering 0.4×0.4 deg between $0.4 < z < 1.5$. We show that the PDF of density contrasts of the VVDS galaxies is an unbiased tracer of the underlying parent distribution up to redshift $z = 1.5$, on scales $R = 8$ and $10 h^{-1}$ Mpc. The second moment of the PDF, i.e., the rms fluctuations of the galaxy density field, is to a good approximation constant over the full redshift baseline investigated: we find that, in redshift space, σ_8 for galaxies brighter than $M_b^* = -20 + 5 \log h$ has a mean value of 0.94 ± 0.07 in the redshift interval $0.7 < z < 1.5$. The third moment, i.e., the skewness, increases with cosmic time: we find that the probability of having underdense regions is greater at $z \sim 0.7$ than it was at $z \sim 1.5$. By comparing the PDF of galaxy density contrasts with the theoretically predicted PDF of mass fluctuations we infer the redshift-, density- and scale-dependence of the biasing function $b(z, \delta, R)$ between galaxy and matter overdensities up to redshift $z = 1.5$. Our results can be summarized as follows: i) the galaxy bias is an increasing function of redshift: evolution is marginal up to $z \sim 0.8$ and more pronounced for $z \gtrsim 0.8$; ii) the formation of bright galaxies is inhibited below a characteristic mass-overdensity threshold whose amplitude increases with redshift and luminosity; iii) the biasing function is non linear in all the redshift bins investigated with non-linear effects of the order of a few to $\sim 10\%$ on scales $> 5 h^{-1}$ Mpc. By subdividing the sample according to galaxy luminosity and colors, we also show that: iv) brighter galaxies are more strongly biased than less luminous ones at every redshift and the dependence of biasing on luminosity at $z \sim 0.8$ is in good agreement with what is observed in the local Universe; v) red objects are systematically more biased than blue objects at all cosmic epochs investigated, but the relative bias between red and blue objects is constant as a function of redshift in the interval $0.7 < z < 1.5$, and its value ($b^{\text{rel}} \sim 1.4$) is similar to what is found at $z \sim 0$.

Key words. cosmology: large-scale structure of Universe – galaxies: distances and redshifts – galaxies: evolution – galaxies: statistics

★ Based on data obtained with the European Southern Observatory Very Large Telescope, Paranal, Chile.

★★ Appendix A is only available in electronic form at <http://www.edpsciences.org>

1. Introduction

The understanding of how matter structures grow via gravitational instability in an expanding Universe is quite well developed and has led to a successful and predictive theoretical framework (e.g., Peebles 1980; Davis et al. 1985).

One of the most critical problems, however, is to understand the complex mechanisms which, on various cosmological scales, regulate the formation and the evolution of luminous structures within the underlying dark-matter distribution. Its solution ultimately relies on the comprehension of the “microscopic” physics which describes how the baryons fall, heat-up, virialize, cool and form stars in the potential wells generated by the dominant mass component of the Universe, i.e., the non-baryonic dark matter (e.g., White & Rees 1978). A zero-order, minimal approach to this investigation consists of “macroscopically” characterizing the cosmological matter fluctuations in terms of a reduced set of fundamental quantities, essentially their positions and mass scales, and in studying how the respective spatial clustering and density amplitudes relate to the corresponding statistics computed for light fluctuations. This comparison scheme is generally referred to as matter-galaxy biasing (e.g., Dekel & Lahav 1999).

An operational definition of bias is conventionally given in terms of continuous density fields by assuming that the local density fluctuation pattern traced by galaxies (δ_g) and mass (δ) are deterministically related via the “linear biasing scheme”

$$\delta_g(\delta) = b\delta \quad (1)$$

where the constant “slope” b is the biasing parameter (Kaiser 1984).

This specific formulation, however, represents a very crude approximation which is not based on any theoretical or physical motivation. It is obvious, for example, that such a model cannot satisfy the physical requirement $\delta_g(-1) = -1$ for any arbitrary value $b \neq 1$. In particular the biasing process could be non local (e.g., Catelan et al. 1998), stochastic (e.g., Dekel & Lahav 1999) and non linear (e.g., Mo & White 1996). Moreover, both theory and numerical simulations predict that the bias grows monotonically from the present cosmic epoch to high redshifts (e.g., Dekel & Rees 1987; Fry 1996; Mo & White 1996; Tegmark & Peebles 1998; Basilakos & Plionis 2001).

From a theoretical perspective, light does not follow the matter distribution on sub-galactic scales, where nearly 90% of dense, low-mass dark matter fluctuations ($M \sim 10^7 - 10^8 h^{-1} M_\odot$) failed to form stars and to become galaxies (e.g., Klypin et al. 1999; Moore et al. 1999; Dalal & Kochanek 2002). A difference in the spatial distribution of visible and dark matter is predicted also on galactic scales, since the radial scaling of density profiles of dark matter halos (Navarro et al. 1997) differs from the three-dimensional radial distributions of light (Sersic and Freeman laws). Galaxy biasing is theoretically expected also on cosmological scales. In particular, simulations of the large-scale structure predict the existence of a difference in the relative distribution of mass and dark halos (e.g., Cen & Ostriker 1992; Bagla 1998; Kravtsov & Klypin 1999) or galaxies (e.g., Evrard, et al. 1994; Blanton et al. 2000; Kayo et al. 2001). Various physical mechanisms for

biasing have been proposed, such as, for example, the peaks-biasing scheme (Kaiser 1984; Bardeen et al. 1986), the probabilistic biasing approach (Coles 1993), or the biasing models derived in the context of the extended Press & Schechter approximation (Mo & White 1996; Matarrese et al. 1998).

Turning to the observational side, the fact that, in the local Universe, galaxies cluster differently according to morphological type (Davis & Geller 1976), surface brightness (Davis & Djorgovski 1985), luminosity (Mauger & Lachièze-Rey 1987), or internal dynamics (White et al. 1988) implies that not all can simultaneously trace the underlying distribution of mass, and that galaxy biasing not only exists, but might also be sensitive to various physical processes. Redshift information that recently became available for large samples of galaxies has significantly contributed to better shaping our current understanding of galaxy biasing, at least in the local Universe. The analysis of the power spectrum (Lahav et al. 2002) and bispectrum (Verde et al. 2002) of the 2dF Galaxy Redshift Survey (2dFGRS Colless et al. 2001) consistently shows that a flux-limited sample of local galaxies ($z < 0.25$), optically selected in the b_J -band ($b_J \leq 19.4$), traces the mass, i.e., it is unbiased, on scales $5 < R(h^{-1}\text{Mpc}) < 30$.

The galaxy correlation function has been measured up to redshift ~ 1 by the CFRS (Le Fèvre et al. 1996), and by the CNOC (Carlberg et al. 2000) surveys giving conflicting evidence on clustering amplitude and bias evolution (see Small et al. 1999). More recently, the analysis of the first season DEEP2 data (Coil et al. 2004) seems to indicate that a combined R -band plus color selected sample is unbiased at $z \sim 1$. On the contrary, measurements of the clustering (Steidel et al. 1998; Giallisco . 1998; Foucaud et al. 2003) or of the amplitude of the count-in-cell fluctuations (Aldeberger et al. 1998) of Lyman-break galaxies (LBGs) at $z \sim 3$ suggest that these objects are more highly biased tracers of the mass density field than are galaxies today. Higher redshift domains have been probed by using photometric redshift information (Arnouts et al. 1999), or compilation of heterogeneous samples (Magliocchetti et al. 2000). Again, the clustering signal appears to come from objects which are highly biased with respect to the underlying distribution of mass.

While there is general observational consensus on the broad picture, i.e., that biasing must decrease with cosmic time, the elucidation of the finer details of this evolution as well as any meaningful comparison with specific theoretical predictions is still far from being secured. Since clustering depends on morphology, color and luminosity, and since most high redshift samples have been selected according to different colors or luminosity criteria, it is not clear, for example, how the very different classes of objects (Ly-break galaxies, extremely red objects or ultraluminous galaxies), which populate different redshift intervals, can be considered a uniform set of mass tracers across different cosmic epochs. Furthermore, the biasing relation is likely to be nontrivial, i.e., non-linear and scale dependent, especially at high redshift (e.g., Somerville et al. 2001).

Only large redshift surveys defined in terms of uniform selection criteria and sampling typical galaxies (or their progenitors), rather than small subclasses of peculiar objects, promise to yield a more coherent picture of biasing evolution.

In particular, the 3D spatial information provided by the VIMOS-VLT Deep Survey (VVDS, Le Fèvre et al. 2005a, hereafter Paper I) should allow us to investigate the mass and scale dependence, as well as to explore the time evolution of the biasing relation between dark matter and galaxies for a homogeneous, flux-limited ($I \leq 24$) sample of optically selected galaxies.

The intent of this paper is to provide a measure, on some characteristic scales R , over the continuous redshift interval $0.4 < z < 1.5$, of the *local, non-linear, deterministic* biasing function

$$b = b(z, \delta, R). \quad (2)$$

The goal is to provide an observational benchmark for theories predicting the efficiency of structure formation, or semi-analytical simulations of galaxy evolution. The problem, however, is to find an optimal strategy to evaluate the biasing function in the quasi-pencil-beam geometry of the first-epoch VVDS survey. At present, the angular size of the first-epoch VVDS redshift cone ($\sim 0.5 \text{ deg}^2$) does not allow us to constrain the biasing function using high order moments of the galaxy distribution (for example the 3-point correlation function). Moreover, we cannot determine the biasing function simply regressing the galaxy fluctuations (δ_g) versus mass fluctuations (δ). The fundamental limitation preventing such an intuitive comparison is evident: it is easy to “pixelize” the survey volume and to measure the galaxy fluctuations in each survey cell, but, since the VVDS is not a matter survey, it is much less straightforward to assign to each cell a mass density value. Thus, in this paper, we take an orthogonal approach and we infer the biasing relation $\delta_g = \delta_g(\delta)$ between mass and galaxy overdensities from their respective probability distribution functions (PDFs) $f(\delta)$ and $g(\delta_g)$: assuming a one-to-one mapping between mass and galaxy overdensity fields, conservation of probability implies

$$\frac{d\delta_g(\delta)}{d\delta} = \frac{f(\delta)}{g(\delta_g)}. \quad (3)$$

The advantage over other methods is that we can explore the functional form of the relationship $\delta_g = b(z, \delta, R)\delta$ over a wide range in mass density contrasts, redshift intervals and smoothing scales R without specifying any a priori parametric functional form for the biasing function.

In pursuing our approach, we assume that the PDF of matter overdensities $f(\delta)$ is satisfactorily described by theory and N-body simulations. What we will try to assess explicitly, is the degree at which the measured PDF of the VVDS overdensities $g(\delta_g)$ reproduces the PDF of the underlying parent population of galaxies. The large size and high redshift sampling rate of the VVDS spectroscopic sample, together with the multi-color information in the B, V, R, I filters of the parent photometric catalog and the relatively simple selection functions of the survey, allow us to check for the presence of observational systematics in the data. In principle, this analysis helps us to constrain the range of the parameter space where first-epoch VVDS data can be analyzed in a statistically unbiased way and results can be meaningfully interpreted.

The outline of the paper is the following: in Sect. 2 we briefly describe the first-epoch VVDS data sample. In Sect. 3 we introduce the technique applied for reconstructing the three-dimensional density field traced by VVDS galaxies, providing details about corrections for various selection effects. In Sect. 4 we outline the construction of the PDF of galaxy overdensities and test its statistical representativity. We then derive the PDF of VVDS density contrasts and analyze its statistical moments. In Sect. 5 we review the theoretical properties of the analogous statistics for mass fluctuations. Particular emphasis is given to the problem of projecting the mass PDF derived in real space into redshift-perturbed comoving coordinates in the high redshift Universe. The method for computing the biasing function is introduced and tested against possible systematics in Sect. 6. VVDS results are presented and discussed in Sect. 7. and compared to theoretical models of biasing evolution in Sect. 8. Conclusions are drawn in Sect. 9.

The coherent cosmological picture emerging from independent observations and analysis motivate us to frame all the results presented in this paper in the context of a Λ CDM cosmological model with $\Omega_m = 0.3$ and $\Omega_\Lambda = 0.7$. Throughout, the Hubble constant is parameterized via $h = H_0/100$. All magnitudes in this paper are in the AB system (Oke & Gunn 1983), and from now on we will drop the suffix AB.

2. The first-epoch VVDS redshift sample

The primary observational goal of the VIMOS-VLT Redshift Survey as well as the survey strategy and first-epoch observations in the VVDS-0226-04 field (from now on simply VVDS-02h) are presented in Paper I.

In order to minimize selection biases, the VVDS survey in the VVDS-02h field has been conceived as a purely flux-limited ($17.5 \leq I \leq 24$) survey, i.e., no target pre-selection according to colors or compactness is implemented. Stars and QSOs have been a posteriori removed from the final redshift sample. Photometric data in this field are complete and free from surface brightness selection effects, up to the limiting magnitude $I = 24$ (Mc Cracken et al. 2003).

First-epoch spectroscopic observations in the VVDS-02h field were carried out using the VIMOS multi-object spectrograph (Le Fèvre et al. 2003) during two runs between October and December 2002 (see Paper I). VIMOS observations have been performed using 1 arcsec wide slits and the LRRed grism which covers the spectral range $5500 < \lambda(\text{\AA}) < 9400$ with an effective spectral resolution $R \sim 227$ at $\lambda = 7500 \text{\AA}$. The accuracy in redshift measurements is $\sim 275 \text{ km s}^{-1}$. Details on observations and data reduction are given in Paper I, and in Le Fèvre et al. (2004).

The first-epoch VVDS-02h data sample extends over a sky area of $0.7 \times 0.7 \text{ deg}$ (which was targeted according to a 1, 2 or 4 passes strategy, i.e., giving to any single galaxy in the field 1, 2 or 4 chances to be targeted by VIMOS masks (see Fig. 12 of Paper I) and has a median depth of about $z \sim 0.76$. It contains 6582 galaxies with secure redshifts (i.e., redshift determined with a quality flag ≥ 2 (see Paper I)) and probes a comoving volume (up to $z = 1.5$) of nearly $1.5 \times 10^6 h^{-3} \text{ Mpc}^3$ in a standard

Λ CDM cosmology. This volume has transversal dimensions $\sim 37 \times 37 h^{-1}\text{Mpc}$ at $z = 1.5$ and extends over $3060 h^{-1}\text{Mpc}$ in radial direction.

For this study we define a sub-sample (VVDS-02h-4) with galaxies having redshift $z < 1.5$ and selected in a continuous sky region of 0.4×0.4 deg which has been homogeneously targeted four times by VIMOS slitmasks. Even if we measure redshifts up to $z \sim 5$ and in a wider area, the conservative angular and redshift limits bracket the range where we can sample in a denser way the underlying galaxy distribution and, thus, minimize biases in the reconstruction of the density field (see the analysis in Sect. 4.1). The VVDS-02h-4 subsample contains 3448 galaxies with secure redshift (3001 with $0.4 < z < 1.5$) and probes one-third of the total VVDS-02h volume. This is the main sample used in this study.

3. The density field reconstruction scheme

The first ingredient we need in order to derive the biasing relation

$$\delta_g = b(z, \delta, R)\delta \quad (4)$$

is a transformation scheme for diluting an intrinsic point-like process, such as the galaxy distribution in a redshift survey, into a continuous 3D overdensity field (see the review by Strauss & Willick 1995). We write the dimensionless density contrast at the *comoving* position \mathbf{r} , smoothed over a typical dimension R as

$$\delta_g(\mathbf{r}, R) = \frac{\rho_g(\mathbf{r}, R) - \bar{\rho}_g}{\bar{\rho}_g}. \quad (5)$$

and we define (e.g., Hudson 1993) the smoothed number density of galaxies above the absolute magnitude threshold \mathcal{M}^c as the convolution between Dirac's delta functions and some arbitrary filter

$$\rho_g(\mathbf{r}, R, < \mathcal{M}^c) = \sum_i \frac{\delta^D(\mathbf{r} - \mathbf{r}_i) * F\left(\frac{|\mathbf{r} - \mathbf{r}_i|}{R}\right)}{S(r_i, \mathcal{M}^c)\Phi_z(m)} \quad (6)$$

where the sum is taken over all the galaxies in the sample, $S(r, \mathcal{M}^c)$ is the distance-dependent selection (or incompleteness) function of the sample (see Sect. 3.2), $\Phi_z(m)$ is the redshift sampling function (see Sect. 3.3) and $F(|\mathbf{r}|/R)$ is a smoothing kernel of width R . In this paper, the smoothing window F is modeled in terms of a normalized Top-Hat (TH) filter

$$F\left(\frac{|\mathbf{r}|}{R}\right) = \frac{3}{(4\pi R^3)}\Theta\left(1 - \frac{|\mathbf{r}|}{R}\right), \quad (7)$$

where Θ is the Heaviside function, defined as $\Theta(x) = 1$ for $0 \leq x \leq 1$, and $\Theta(x) = 0$ elsewhere.

Within this weighting scheme, shot-noise errors are evaluated by computing the variance of the galaxy field

$$\epsilon(\mathbf{r}) = \frac{1}{\bar{\rho}_g} \left[\sum_i \left(\frac{F\left(\frac{|\mathbf{r} - \mathbf{r}_i|}{R}\right)}{S(r_i, \mathcal{M}^c)\Phi_z(m_i)} \right)^2 \right]^{1/2}. \quad (8)$$

Note that all coordinates are comoving and that the mean density $\bar{\rho}_g$ depends on cosmic time. Since we observe an evolution of nearly a factor of two in the mean density of galaxies

brighter than $\mathcal{M}_B^*(z = 0)$ from redshift $z = 0$ up to $z = 1$ (see Ilbert et al. 2005, hereafter Paper II), we compute the value of $\bar{\rho}_g$ at position \mathbf{r} (corresponding to some look-back time t) with Eq. (6) by simply averaging the galaxy distribution in survey slices $\mathbf{r} \pm R_s$ where $R_s = 400 h^{-1}\text{Mpc}$. We note that conclusions drawn in this paper depend very weakly on the choice of R_s in the interval $200 < R_s(h^{-1}\text{Mpc}) < 600$.

In this paper, the density field is evaluated at positions \mathbf{r} in the VVDS-02h volume that can be either random or regularly displaced on a 3D grid (see Sect. 3.4). Even if we correct for the different sampling rate in the VVDS-02h field (by weighting each galaxy by the inverse of redshift sampling function $\Phi_z(m)$) we always select, for the purposes of our analysis, only the density fluctuations recovered in spheres having at least 70% of their volume in the denser 4-passes volume. This in order to minimize the Poissonian noise due to the sparser redshift sampling outside the VVDS-02h-4 field.

We also consider only volumes above the redshift threshold z_t where the transversal dimension L of the first-epoch VVDS-02h cone is $L(z_t) > 2R$. As an example, for TH windows of size $R = 5(10)h^{-1}\text{Mpc}$ we have $z_t \sim 0.4(0.7)$. Within the redshift range $0.4 < z < 1.5$ the VVDS-02h field contains 5252 galaxies (of which 3001 are in the four passes region).

Note that we have characterized the galaxy-fluctuation field in terms of the number density contrast, instead of the luminosity density contrast, because the former quantity is expected to show a time-dependent variation which is more sensitive to the galaxy evolution history (formation and merger rates, for example). Moreover, as described in Sect. 3.4, a robust description of the density field and a reliable determination of the PDF shape can be obtained only minimizing the shot noise component of the scatter; this is more easily done by considering galaxy number densities rather than galaxy light densities.

The most critical elements of the smoothing process are directly readable in Eq. (6): we must first evaluate galaxy absolute magnitudes at each redshift in the most reliable way, then specify the selection function and the redshift sampling rate of the VVDS survey. In the next sections we will describe how these quantities have been evaluated.

3.1. The K -correction

The absolute magnitude is defined as:

$$M^r = m^o - 5 \log d_L(z, \Omega) - K(z, \text{SED}) \quad (9)$$

where the suffixes r and o designate respectively the rest-frame band in which the absolute magnitude is computed and the band where the apparent magnitude is observationally measured, and d_L is the luminosity distance evaluated in a given a priori cosmology (i.e., using an appropriate set of cosmological parameters $\Omega \equiv [\Omega_m, \Omega_\Lambda]$).

The correction factor K , which depends on redshift and the spectral energy distribution (SED), accounts for the fact that the system response in the observed frame corresponds to a narrower, bluer rest-frame passband, depending on the redshift of the observed object. A complete description of the application of this transformation technique to VVDS galaxies is detailed in Paper II.

The estimate of the galaxy absolute luminosity is thus affected by the uncertainties introduced by probing redshift regimes where the k -correction term cannot be neglected. Using mock catalogs simulating the VVDS survey, we have shown (see Fig. A.1 in Paper II) that the errors in the recovered absolute magnitude are significantly smaller in the B band ($\sigma_B = 0.08$) than in the I band ($\sigma_I = 0.17$). Thus, in what follows we will use absolute luminosities determined in the B -band rest-frame.

3.2. The radial selection function

Since our sample is limited at bright and faint apparent magnitudes ($17.5 \leq I \leq 24$), at any given redshift we can only observe galaxies in a specific, redshift-dependent, absolute magnitude range. It is usual to describe the sample radial incompleteness by defining the selection function in terms of the galaxy luminosity function $\varphi(M)$

$$S(r, M^c) = \frac{\int_{M_b(r)}^{M_f(r)} \varphi(M) dM}{\int_{-\infty}^{M^c} \varphi(M) dM}. \quad (10)$$

Here $M_b(r)$ and $M_f(r)$ are the B -band absolute magnitudes which correspond, at distance r , to the I -band limiting apparent magnitudes $m_b = 17.5$ and $m_f = 24$ respectively (see discussion at the end of the section). Since Eq. (9) also depends on galaxy colors, we compute its mean value at distance r by a weighted average over the population mix observed at distance r .

The VVDS luminosity function (LF) has been derived in Paper II and is characterized by a substantial degree of evolution over the redshift range $0 < z < 1.5$. Therefore, we estimate $\varphi(M)$ in the B band at any given position in the redshift interval $[0, 1.5]$ by interpolating, with a low order polynomial function, the Schechter shape parameters α and M^* given in Table 1 of Paper II.

Assuming $M_B^c = -15$ in Eq. (10), which corresponds to the limiting absolute magnitude over which the LF of the VVDS-02h sample can be robustly constrained in the lowest redshift bins, the selection function exponentially falls by nearly 2 orders of magnitude in the redshift range up to $z < 1.5$. Thus, the density field reconstruction strongly depends on the radial selection function used especially at high redshifts, where Eq. (10) can be affected by possible systematics in the determination of the LF or in the measurements of faint magnitudes. Therefore, we will also analyze volume-limited sub-samples, which are essentially free from these systematics.

Since in a magnitude limited survey progressively brighter galaxies are selected as a function of redshift, a volume-limited sample also allows us to disentangle spurious luminosity-dependent effects from the measurement of the redshift evolution of the biasing function.

3.3. The redshift sampling rate

As for most redshift surveys, the VVDS does not target spectroscopically all the galaxies that satisfy the given flux limit criteria in the selected field of view (see Paper I). Because of

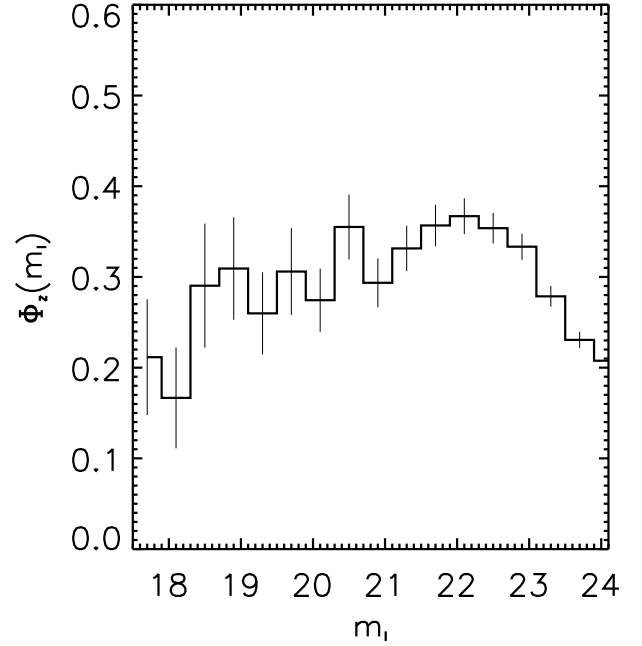


Fig. 1. The VVDS redshift sampling rate in the four-passes VVDS-02h-4 region is plotted versus the observed apparent magnitude in the I -band. The mean redshift sampling rate is ~ 0.3 .

the sparse sampling strategy, we have to correct the density estimator with a sampling rate weight Φ_z in order to reconstruct the real underlying galaxy density field in a statistically unbiased way.

The VVDS redshift sampling rate is the combination of two effects: i) only a fraction of the galaxies ($\sim 40\%$, see Paper I) in the photometric sample is targeted (*target sampling rate*); ii) and only a fraction of the targeted objects ($\sim 80\%$ see Paper I) yield a redshift (*spectroscopic success rate*). We can model this correction term, by assuming, to a first approximation, that the sampling rate depends only on the apparent magnitude. Since the VVDS targeting strategy is optimized to maximize the number of slits on the sky, the selection of faint objects is systematically favored. Inversely, the ability of measuring a redshift degrades progressively towards fainter magnitudes, i.e., for spectra having lower signal-to-noise ratios (the spectroscopic success rate decreases from $>90\%$ at $I \sim 22$ down to $\sim 60\%$ at $I \sim 24$). These two opposite effects conspire to produce the magnitude-dependent sampling rate function shown in Fig. 1. Clearly, with such an approximation, we neglect any possible dependence of the sampling rate from other important parameters such as for example surface brightness, spectral type or redshift. However, in Sect. 3.2 of Paper II we showed, using photometric redshifts, that any systematic sampling bias introduced by a possible redshift dependence of the spectroscopic success rate is expected to affect only the tails of our observed redshift distribution ($z < 0.5$ and $z > 1.5$) i.e., redshift intervals not considered in this study (see Sect. 3).

We describe the VVDS sampling completeness $\Phi_z(m)$ at a given magnitude m , as the fraction of objects with measured

redshifts N_z over the number N of objects detected in the photometric catalog.

$$\Phi_z(m) = \frac{\sum_{i=1}^{N_z} w(m - m_i, dm)}{\sum_{i=1}^N w(m - m_i, dm)} \quad (11)$$

with the window function w defined as:

$$w(m - m_i, dm) = \begin{cases} 1 & \text{if } |m - m_i| < dm/2 \\ 0 & \text{otherwise.} \end{cases} \quad (12)$$

Working in low resolution mode (i.e. allowing up to 4 galaxies to have spectra aligned along the same dispersion direction with a typical sky separation of 2 arcmin) and observing the same region of sky 4 times for a total of about 16 hours (four-passes strategy), we can measure redshifts for nearly 30% of the parent photometric population, as shown in Fig. 1. In other terms, on average, nearly one over three galaxies with magnitude $I \leq 24$ has a measured redshift in the four-passes VVDS region. This high spatial sampling rate is a critical factor for minimizing biases in the reconstruction of the 3D density field of galaxies. In particular we note that our $I \leq 24$, $z < 1.5$ VVDS-02h-4 sample is characterized by an effective mean inter-particle separation in the redshift range $[0, 1.5]$ ($\langle r \rangle \sim 5.1 h^{-1}\text{Mpc}$) which is smaller than that of the original CFA sample ($\langle r \rangle \sim 5.5 h^{-1}\text{Mpc}$) used by Davis & Huchra (1981) to reconstruct the 3D density field of the local Universe (i.e., out to $\sim 80 h^{-1}\text{Mpc}$). At the median depth of the survey, i.e., in the redshift interval 0.75 ± 0.05 , the mean inter-particle separation is $4.4 h^{-1}\text{Mpc}$, a value nearly equal to the mean inter-particle separation at the median depth of the 2dFGRS. We finally note that in the redshift range $[0.7, 1.35]$, which is also covered by the DEEP2 survey, the VVDS mean-inter-particle separation is $5.5 h^{-1}\text{Mpc}$ compared to the value $\sim 6.5 h^{-1}\text{Mpc}$ inferred from the values quoted by Coil et al. (2004) for their most complete field, currently covering 0.32 deg^2 .

By dividing the VVDS-02h-4 field in smaller cells and repeating the analysis, we conclude that the sampling rate does not show appreciable variations, i.e., the angular selection function can be considered constant for the purposes of our analysis. This corresponds to the fact that the success rate in redshift measurement in each VIMOS quadrant (i.e., the *spectroscopic success rate* per mask) is, to a good approximation, constant and equal to $\sim 80\%$ (see Paper I).

3.4. Shot noise

In a flux-limited sample, the shot noise in the density field is an increasing function of distance (see Eq. (8)). One may correct for the increase of the mean VVDS inter-particle separation as a function of redshift (and thus the increase of the variance of the density field) by opportunely increasing the length of the smoothing window (e.g., Strauss & Willick 1995). However, since we are interested in comparing the fluctuations recovered on the same scale at different redshifts in a flux limited survey, we take into account the decreased sampling sensitivity of the survey at high redshift in an alternative way.

We deconvolve the signature of this noise from the density maps by applying the Wiener filter technique

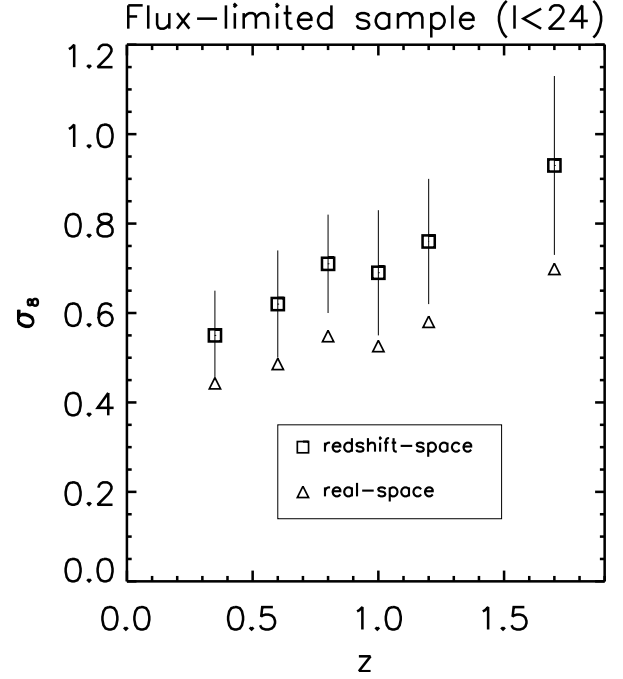


Fig. 2. The real- and redshift-space rms fluctuations of the flux-limited VVDS sample recovered using Eq. (13) and the results of the correlation function analysis presented in Paper III are plotted at six different redshifts in the interval $0.4 < z < 1.7$.

(cf. Press et al. 1992) which provides the *minimum variance* reconstruction of the smoothed density field, given the map of the noise and the a priori knowledge of the underlying power spectrum (e.g., Lahav et al. 1994). The application of the Wiener denoising procedure to the specific geometry of the VVDS sample is described in detail in Appendix A.

Here we note that the Wiener filter requires a model for the underlying 3D power spectrum $P(k, z)$ which we compute, over the frequency scales where the correlation function of VVDS galaxies is well constrained ($0.06 \leq k \leq 10$), as (see Eq. (48) in Appendix A):

$$P(k, z) = 4\pi \frac{r_0(z)^{\gamma(z)}}{k^{3-\gamma(z)}} \Gamma(2 - \gamma(z)) \sin \frac{\gamma(z)\pi}{2}, \quad (13)$$

where the normalization $r_0(z)$ and the slope $\gamma(z)$ of the correlation function at redshift z have been derived by interpolating the values measured in various redshift intervals of the VVDS-02h volume by le Fèvre et al. (2005, hereafter Paper III).

The variance of the galaxy distribution on a $8 h^{-1}\text{Mpc}$ scale in the VVDS-02h sample can be obtained by integrating Eq. (13) using a TH window of radius $8 h^{-1}\text{Mpc}$ and the (r_0, γ) parameters of the VVDS correlation function

$$\sigma_8^2 = \frac{1}{2\pi^2} \int P(k) \tilde{F}_k^2 k^2 dk \quad (14)$$

where \tilde{F}_k is the Fourier transform of the TH filter (see Eq. (45) in Appendix A). Note that, by integrating the power spectrum down to $k \rightarrow 0$, i.e., extrapolating the power law shape of Eq. (13) beyond $L \sim 100 h^{-1}\text{Mpc}$ ($k \lesssim 0.06$), one would revise upwards the value of σ_8 by $\sim 2\%$ (at $z = 0.35$) and by $\sim 4\%$ at $z = 1.4$. Since, however, the amplitude of the power spectrum

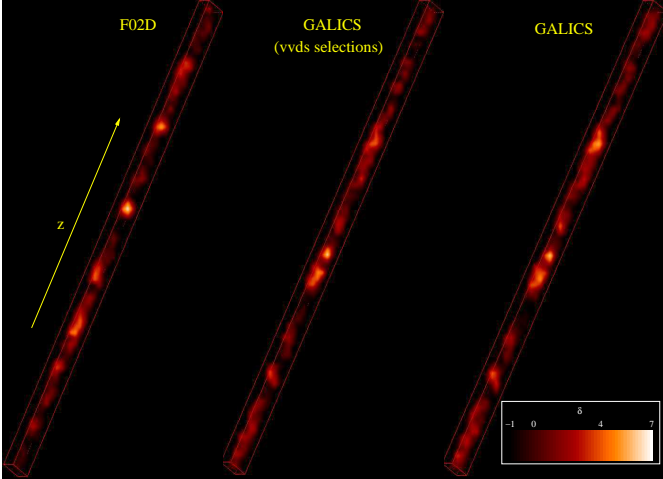


Fig. 3. 3D density field traced by the galaxy distribution in the VVDS-02h Field (*left*, 1641 galaxies), in the flux-limited ($I \leq 24$) GALICS simulation (*right*, 9450 galaxies), and in the flux-limited GALICS sample after applying the VVDS target selection criteria (*center*, 1656 galaxies). Data span the redshift interval $[0.8, 1.1]$. In each cone, the galaxy distribution is continuously smoothed using a TH window function with $R = 5 h^{-1}\text{Mpc}$ which nearly corresponds to the mean inter-particle separation in this same redshift interval. The cone metric was computed assuming a ΛCDM cosmology and the correct axis ratio between transversal and radial dimensions has been preserved. The cones have approximate transverse dimensions of $28 h^{-1}\text{Mpc}$ at $z = 1$ and extend over $527 h^{-1}\text{Mpc}$.

on large scales is expected to downturn and to be systematically lower than predicted by Eq. (13), we safely conclude that, with our computation scheme, the inferred σ_8 value should be biased low by no more than $\sim 2\%$ and 4% in the first and last redshift bin, respectively.

Projecting the results from real-space into the redshift-distorted space (see Sect. 5), i.e., implementing the effects of large-scale streaming motions, we infer that the rms of galaxy fluctuations are $\sigma_8 = [0.55 \pm 0.10, 0.62 \pm 0.12, 0.71 \pm 0.11, 0.69 \pm 0.14, 0.75 \pm 0.14, 0.92 \pm 0.20]$ at redshift $z = [0.35, 0.6, 0.8, 1.2, 1.2, 1.65]$ (see Fig. 2). Thus, the amplitude of σ_8 for a flux-limited $I \leq 24$ sample increases as a function of redshift by nearly 70% between $z \sim 0.3$ and $z \sim 1.7$.

3.5. The 3-dimensional VVDS density field

The VVDS-02h galaxy density field reconstructed on a scale $R = 5 h^{-1}\text{Mpc}$ and in the redshift bin $0.8 < z < 1.1$ is visually displayed in the left most panel of Fig. 3. Note that the chosen smoothing scale nearly corresponds to the mean inter-galaxy separation in the VVDS-02h sample at $z \sim 1$. Figure 3 shows the regular patterns traced by over- and under-dense regions in the selected redshift interval. Specifically, we note that, in this redshift slice, there are over- and under-dense regions which extend over characteristic scales as large as $\sim 100 h^{-1}\text{Mpc}$. A more complete discussion of the “cartography” in such deep regions of the Universe is presented by Le Fèvre et al. (2005c).

In the same figure, we also display the density field reconstructed in an analogous redshift range, using the GALICS

simulation (Galaxies in Cosmological Simulations, Hatton et al. 2003). GALICS combines cosmological simulations of dark matter with semi-analytic prescriptions for galaxy formation to produce a fully realistic deep galaxy sample. In particular we plot the density field of the $I \leq 24$ flux-limited simulation as well as the density field recovered after applying to the pure flux-limited simulation all the VVDS target selection criteria (see Sect. 4.1). No qualitative difference between the density fields reconstructed before and after applying to the simulation all the survey systematics is seen.

Clearly, a more quantitative assessment of the robustness and reliability of the VVDS overdensity field can be done by studying its PDF.

4. The PDF of galaxy fluctuations

Once the three-dimensional field of galaxy density contrasts δ_g has been reconstructed on a given scale R , one can fully describe its properties by deriving the associated PDF $g(\delta_g)$. This statistical quantity represents the normalized probability of having a density fluctuation in the range $(\delta_g, \delta_g + d\delta_g)$ within a region of characteristic length R randomly located in the survey volume.

While the shape of the PDF of mass fluctuations at any given cosmic epoch is theoretically well constrained from CDM simulations (see next section), little is known about the observational PDF of the general population of galaxies in the high redshift Universe. Even locally, this fundamental statistics has been often overlooked (but see Ostriker et al. 2003). Notwithstanding, the shape of the galaxy PDF is strongly sensitive to the effects of gravitational instability and galaxy biasing, and its redshift dependence encodes valuable information about the origin and evolution of galaxy density fluctuations.

The shape of the PDF can be characterized in terms of its statistical moments. In particular the variance of a zero-mean field (such as the overdensity field we are considering) is

$$\sigma_R^2 = \langle \delta_g^2 \rangle_R \equiv \int_0^\infty \delta_g^2 g_R(\delta_g) d\delta_g. \quad (15)$$

Higher moments can be straightforwardly derived (see Bernardeau et al. 2002, for a review). In the following, we will drop the suffix R , unless we need to emphasize it.

4.1. Estimating reconstruction systematics using mock catalogs

For the purposes of our analysis, it is imperative to check that the various instrumental selection effects as well as the VVDS observing strategy are not compromising the determination of the PDF of galaxy fluctuations. In this section, we explore the region of the parameter space (essentially redshift and smoothing scales) where the PDF of VVDS overdensities traces in a statistically unbiased way the underlying parent distribution.

Possible systematics can be hidden in the reconstructed PDF essentially because i) the VVDS redshift sampling rate is not unity; ii) the slitlets are allocated on the VIMOS masks with different constraints along the dispersion and the spatial

axis, and iii) the VIMOS field of view is splitted in four different rectangular quadrants separated by a vignetting cross.

We have addressed point iii) by designing a specific telescope pointing strategy which allows us to cover in a uniform way the survey sky region (see the telescope pointing strategy shown in Fig. 1 of Paper I). With the adopted survey strategy, we give to each galaxy in the VVDS-02h-4 field four chances to be targeted by VIMOS, thus increasing the survey sampling rate (nearly 1 galaxy over 3 with magnitude $I \leq 24$ has a measured redshift).

Concerns about points i) and ii) can be directly addressed using galaxy simulations covering a cosmological volume comparable to the VVDS one. Thanks to the implementation of the Mock Map Facility (MoMaF, Blaizot et al. 2003), it is possible to convert the GALICS 3D mocks catalog into 2D sky images, and handle the 2D projection of the simulation as a pseudo-real imaging survey. Pollo et al. (2005), have then built a set of 50 fully realistic mock VVDs surveys from the GALICS simulations to which the whole observational pipeline and biases has been applied. By comparing specific properties of the resulting pseudo-VVDS sample with the true underlying properties of the pseudo-real Universe from which the sample is extracted, we can directly explore the robustness, as well as the limits, of the particular statistical quantities we are interested in.

In brief these include adding to the 3D galaxy mocks a randomly simulated distribution of stars to mimic the same star contamination affecting our survey. Next, we have masked the sky mocks using the VVDS photometric masks, i.e., we have implemented the same geometrical pattern of excluded regions with which we avoid to survey sky regions contaminated by the presence of bright stars or photometric defects. Then, we have extracted the spectroscopic targets by applying the target selection code (VMMPS, Bottini et al. 2005) to the simulated 2D sky distribution. To each GALICS redshift, which incorporates the Doppler contribution due to galaxy peculiar velocities, we have added a random component to take into account errors in z measurements. Finally, we have processed the selected objects implementing the same magnitude-distribution of failures in redshift measurements which characterizes the first-epoch observations of the VVDS survey (see Paper I and Fig. 1).

Since GALICS galaxies have magnitudes simulated in the same 4 bands surveyed by VVDS (B, V, R, I), we have applied the K-correction to obtain rest frame absolute magnitudes and we have empirically re-derived all the selection functions for the mock catalogs according to the scheme presented in Sect. 3. In this way we can also check the robustness of the techniques we apply for computing absolute magnitudes and selection functions (see Paper II).

The PDF of galaxy overdensities obtained from the mock samples has been finally compared to the PDF of the parent population. For brevity, in what follows, we will call *s-samples* (survey-samples) the mocks simulating the VVDS redshift survey and *p-sample* (parent sample) the whole GALICS simulation flux-limited at $17.5 \leq I \leq 24$.

The density contrasts have been calculated as described in Sect. 3. In the following, we will restrict our analysis to the set of smoothing scales in the interval $R = (5, 10) h^{-1}\text{Mpc}$. The

choice of these particular limits is motivated by the fact that $5 h^{-1}\text{Mpc}$ is the minimum smoothing scale for which the reconstructed density field is unbiased over a substantial redshift interval (see discussion below). Note, also, that below this typical scale, linear regimes approximations, largely used in this paper, do not hold anymore. The upper boundary is constrained by the transverse comoving dimensions covered by the first-epoch VVDS data (see Sect. 2), which is still too small for being partitioned using bigger scale-lengths without introducing significant noise in the reconstructed PDF (see the transverse comoving dimension of the VVDS-02h field quoted in Sect. 2).

The PDF of the galaxy density contrasts computed using the *s*-sample is compared to the parent distribution inferred from the *p*-sample in Fig. 4. We conclude that the distribution of galaxy overdensities of the *s*-samples for $R = 8$ and $10 h^{-1}\text{Mpc}$ scales is not biased with respect to the underlying distribution of *p*-sample galaxy fluctuations. Thus, the VVDS density field reconstructed on these scales is not affected by the specific VVDS observational strategy.

It is evident in Fig. 4 that the VVDS redshift sampling rate is not high enough to map in an unbiased way the low density regions of the Universe ($\log(1 + \delta_g) \lesssim -0.5$) when the galaxy distribution is smoothed on scales of $5 h^{-1}\text{Mpc}$. Figure 5 shows that incompleteness in underdense regions is a function of redshift, with the bias in the low-density tail of the PDF developing and increasing as the redshift increases. As a rule of thumb, the PDF of the *s*-sample starts to deviate significantly from the parent PDF when the mean inter-galactic separation $\langle r \rangle$ of the survey sample is larger than the scale R on which the field is reconstructed. Since we measure $\langle r(z=1) \rangle \sim 5 h^{-1}\text{Mpc}$ ($\langle r(z=1.5) \rangle \sim 8 h^{-1}\text{Mpc}$), the PDF of the density field recovered using a TH filter of radius $5 h^{-1}\text{Mpc}$ is effectively unbiased (at least over the density range we are interested in, i.e., $\log(1 + \delta) > -1$) only if the sample is limited at $z \lesssim 1$. Therefore, in the following, results obtained for $R = 5 h^{-1}\text{Mpc}$ are quoted only up to $z = 1$.

On scales $R > 8 h^{-1}\text{Mpc}$, the agreement between the PDFs of *s*- and *p*-samples holds true also for volume-limited subsamples. Specifically, the 2nd and 3rd moments of the PDF of overdensities recovered using volume-limited *s-samples* on these scales are within 1σ of the corresponding values computed for the parent, volume-limited, *p-samples* in each redshift interval of interest up to $z = 1.5$.

To summarize, the results of simulated VVDS observations presented in this section show that, at least on scales $R \geq 8 h^{-1}\text{Mpc}$, the VVDS PDF describes in an unbiased way the general distribution properties of a sample of $I = 24$ flux-limited galaxies up to redshift 1.5. In other terms, the VVDS density field sampled in this way is essentially free from selection systematics in both low- and high-density regions, and can be meaningfully used to infer the physical bias in the distribution between galaxy and matter. Obviously, the representativeness of the measured PDF of VVDS overdensities with respect to the “universal” PDF up to $z = 1.5$ is a different question. Since the volume probed is still restricted to a limited region of space in one field, the shape and moments of the galaxy PDF derived from the VVDS-02h are expected

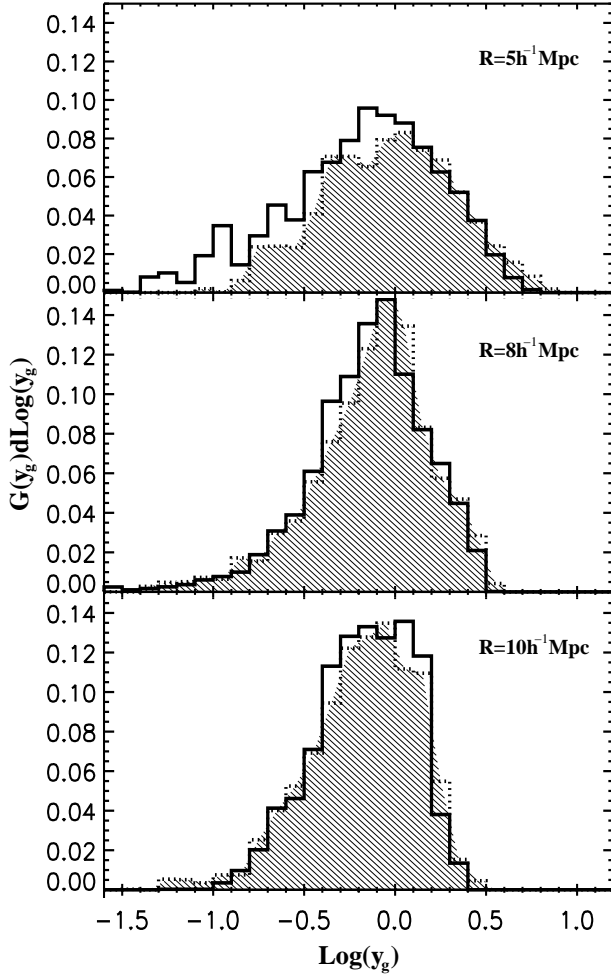


Fig. 4. Tests of the PDF reconstruction scheme using the mocks VVDS samples extracted from GALICS. The differential ($G(y_g)$) probability distribution functions of $y_g = 1 + \delta_g$ for the “observed” s-sample (dotted line, shadowed histogram), and for the parent p-sample (solid-line). Note that the plotted histograms actually corresponds to $G(y_g) = \ln(10)y_g g(y_g)$ since the binning is done in $\log(y_g)$. The logarithmic PDFs are computed for density fields smoothed using TH filters of different sizes (indicated on the top of each panel).

to deviate from the “universal” PDF of galaxies at this redshift because of cosmic variance. Reducing the cosmic variance is one of the main goals of extending the VVDS to 4 independent fields. Anyway, our 50 mock realisations of the VVDS-02h sample allow us to estimate realistic errors that include the contribution from cosmic variance.

4.2. The PDF of VVDS galaxies: results

Let us then investigate the evolution, as a function of the look-back time, of the observed PDF of VVDS galaxy fluctuations.

In an apparent magnitude-limited survey such as the VVDS, only brighter galaxies populate the most distant redshift bins, whereas fainter galaxies are visible only at lower redshifts. As more luminous galaxies tend to cluster more strongly than fainter ones (e.g., Hamilton 1988; Croton et al. 2004), the

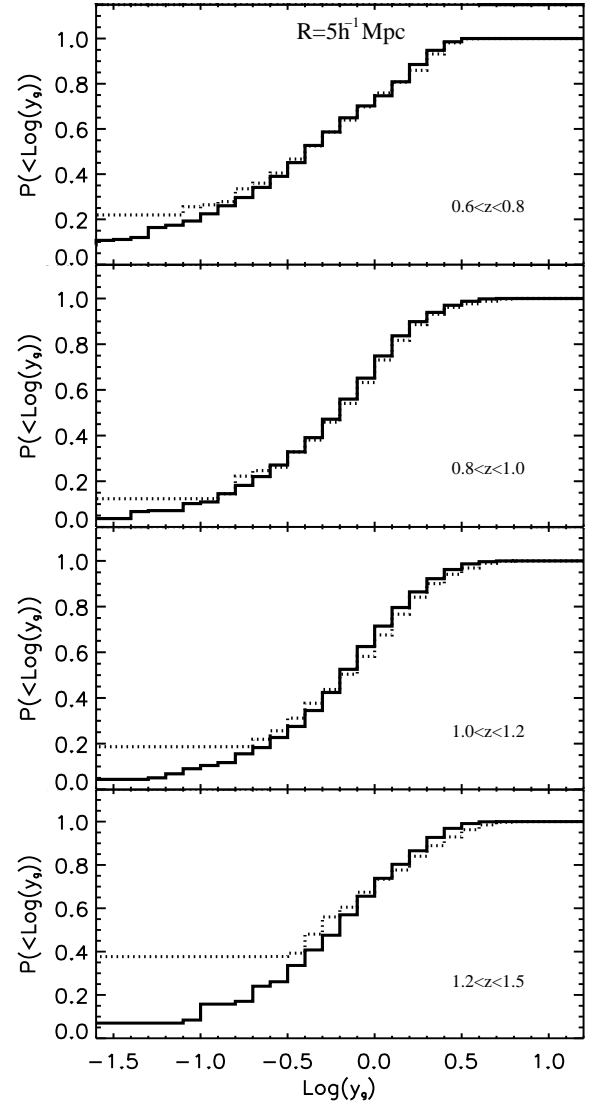


Fig. 5. The cumulative distribution function of density contrasts $y_g = 1 + \delta_g$, on scales $R = 5 h^{-1} \text{Mpc}$, as recovered in different redshift intervals for the s-sample (dotted line), and for the parent p-sample (solid-line). The flattened pedestal at the low-density end of the cumulative distribution is due to low density regions in the p-sample that are spuriously sampled as empty regions ($\delta_g = -1$) in the s-sample.

PDF of galaxy fluctuations is expected to be systematically biased as a function of redshift.

This effect is clearly seen in the first correlation analysis of the VVDS (Paper III), and can be minimized by selecting a volume-limited sample. Therefore, we have defined a subsample with absolute magnitude brighter than $M_B^c = -20 + 5 \log h$ in the rest frame B band (~ 1350 galaxies with $0.7 < z < 1.5$ in the VVDS-02h field, ~ 800 of which are in the VVDS-02h-4 region).

This threshold corresponds to the faintest galaxy luminosity visible at redshift $z = 1.5$ in a $I = 24$ flux-limited survey and it is roughly 0.6 magnitudes brighter(fainter) than the value of M_B^c recovered at $z \sim 0(\sim 1.5)$ using the VVDS data (see Paper II). The median absolute magnitude for this volume-limited sample is ~ -20.4 .

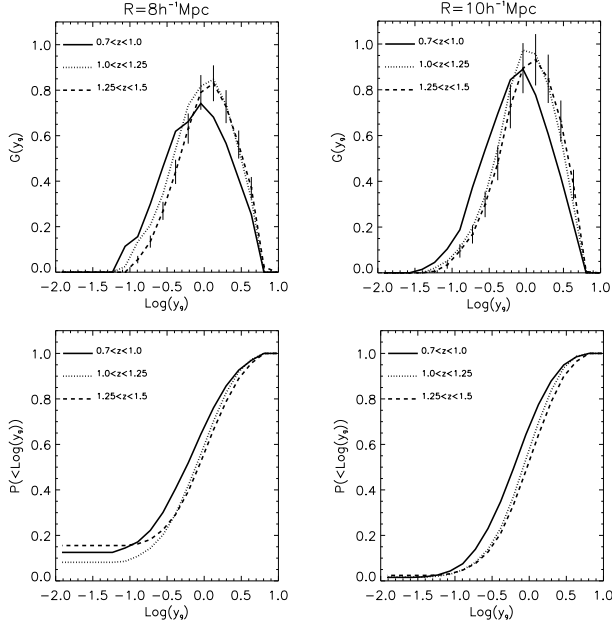


Fig. 6. *Top:* the PDF per units of galaxy overdensities ($G(y) = \ln(10)yg(y)$) is plotted for the volume limited VVDS sample ($M_B = -20 + 5 \log h$) at three different redshifts. The PDFs are computed for density fields recovered using TH filters of size 8 and $10 h^{-1} \text{Mpc}$. *Bottom:* the corresponding cumulative distributions. Errorbars (which, for clarity, are plotted only for the redshift bin $1.25 < z < 1.5$) represent the Poissonian uncertainties. The flattened pedestal at the low-density end of the cumulative distribution is contributed by empty regions ($\delta_g = -1$).

We note, however, that the populations of galaxies with the same luminosity at different redshifts may actually be different. As we have shown in Paper II, we measure a substantial degree of evolution in the luminosity of galaxies, and, as a consequence, with our absolute magnitude cut we are selecting $M^* + 0.6$ galaxies at $z = 1.5$, but $M^* - 0.6$ galaxies at $z = 0$. Thus, the clustering signal at progressively earlier epochs may not be contributed by the progenitors of the galaxies that are sampled at later times in the same luminosity interval.

The PDF of density fluctuations, in various redshift intervals, and traced on scales of 8 and $10 h^{-1} \text{Mpc}$ by VVDS galaxies brighter than M_B^* , is presented in Fig. 6. Note that the analysis of the previous section guarantees that, on these scales, the VVDS PDF fairly represents the PDF of the real underlying population of galaxies up to $z = 1.5$. Figure 6 shows how the shape of the measured galaxy PDF changes across different cosmic epochs.

A Kolmogorov-Smirnov test confirms that the PDFs at different cosmic epochs are statistically different (i.e., the null hypothesis that the three distributions are drawn from the same parent distribution is rejected with a confidence $P_{\text{KS}} > 1 - 10^{-6}$).

In particular the peak of the PDF in the lowest redshift interval is shifted towards smaller values of the density contrast δ_g when compared to the peak of the PDF in the highest redshift bin. Moreover, the shape of the distribution, also shows a systematic “deformation”. Specifically, we observe the development of a low- δ tail in the PDF as a function of time on both scales investigated. In other terms the probability of having low

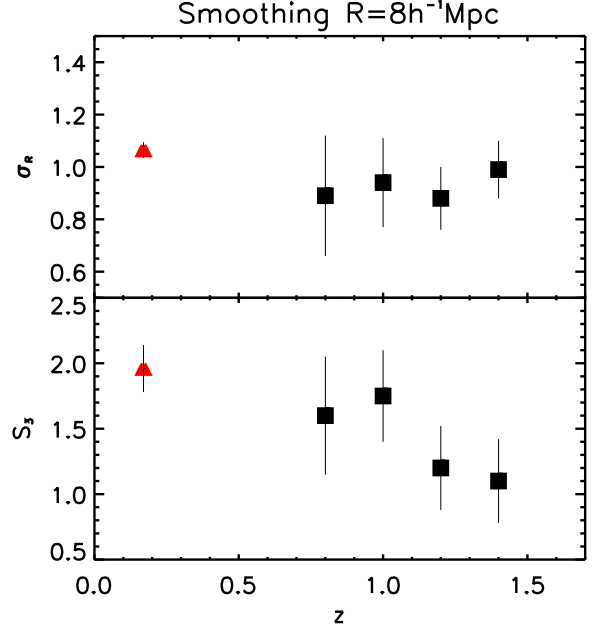


Fig. 7. Redshift evolution of the standard deviation (*upper panel*) and of the skewness (*lower panel*) of the galaxy PDF on a scale $R = 8 h^{-1} \text{Mpc}$ for galaxies brighter than $M = -20 + 5 \log h$. The corresponding local values, estimated on the same scale by Croton et al. (2004) using a subsample of the 2dFGRS having nearly the same median absolute luminosity of our sample, are represented with triangles. Error bars represent 1σ errors, and, in the case of VVDS measurements, include the contribution from cosmic variance. The errorbar on σ_8 of the 2dFGRS is smaller than the symbol size.

density regions increases as a function of cosmic time. For example, underdense regions, defined as the regions where the galaxy density field is $\log(1 + \delta_g) < -0.5$ on a $R = 8 h^{-1} \text{Mpc}$ scale, occupy a fraction of nearly 35% of the VVDS volume at redshift $0.7 < z < 1.0$, but only a fraction of about 25% at redshift $1.25 < z < 1.5$. Similar trends are observed when lowering the absolute luminosity threshold of the volume limited sample (and consequently lowering the upper limit of the redshift interval probed) or when modifying the binning in redshift space.

If galaxies are faithful and unbiased tracers of the underlying dark matter field, then both this effects, the peak shift and the development of a low density tail may be qualitatively interpreted as a direct supporting evidence for the paradigm of the evolution of gravitational clustering in an expanding Universe. At variance with overdense regions which collapse, a net density deficit ($\delta < 0$) in an expanding Universe brings about a sign reversal of the effective gravitational force: a density depression is a region that induces an effective repulsive peculiar gravity (Peebles 1980). If gravity is the engine which drives clustering in an expanding Universe, we thus expect that, as time goes by, low density regions propagate outwards and a progressively higher portion of the cosmological volume becomes underdense.

The observed evolution in the PDF could also indicate the presence of a time-dependent biasing between matter and galaxies. As a matter of fact, it can be easily shown that a monotonic bias, increasing with redshift, offers a natural mechanism

to re-map the galaxy PDF into progressively higher intervals of density contrasts.

We can better discriminate the physical origin of the observed trends, i.e., if they are purely induced by gravitation or strengthened by the collateral and cooperative action of biasing, by studying the evolution of the PDF moments. In Fig. 7 the redshift evolution of the rms (σ) and skewness (S_3) of the overdensity fields (see Table 1) for the $M_B < -20 + 5 \log h$ sample are shown and compared to local measurements.

Following the standard convention within the hierarchical clustering model, we define the skewness S_3 in terms of the volume-averaged two- and three-point correlation functions ($S_3 \equiv \bar{\xi}_3/\bar{\xi}_2^2$) noting that in the case of a continuous δ_g -field with zero mean this expression reduces to $S_3 \equiv \langle \delta_g^3 \rangle / \langle \delta_g^2 \rangle^2$. We do not derive the moments $\langle \delta_g^2 \rangle$ and $\langle \delta_g^3 \rangle$ of the PDF by directly applying the computation scheme given in Eq. (15), but by correcting the count-in-cells statistics for discreteness effects using the Poissonian shot-noise model (e.g., Peebles 1980; Fry 1985, cf. Eqs. (374) and (375) of Bernardeau et al. 2002, possible biases introduced by this estimation technique are discussed by Hui & Gaztañaga 1999). The corresponding values of σ and S_3 for the local Universe (in redshift space) have been derived by Croton et al. 2004 using the 2dFGRS. Here we plot the values corresponding to their $-21 < M - 5 \log h < -20$ subsample, which actually brackets the median luminosity of our volume limited sample.

We can see that the rms amplitude of fluctuations of the VVDS density field, on scales $8 h^{-1} \text{Mpc}$, is into a good approximation constant over the full redshift baseline investigated, with a mean value of 0.94 ± 0.07 over $0.7 < z < 1.5$.

While the strength of clustering of galaxies brighter than $M < -20 + 5 \log h$ does not change much in this redshift interval, each VVDS measurement is lower than the value inferred at $z \sim 0$ by Croton et al. (2004). In particular, our mean value is $\sim 10\%$ smaller than the 2dFGRS value and the difference is significant at $\sim 2\sigma$ level.

The skewness S_3 , which measures the tendency of gravitational clustering to create asymmetries between underdense and overdense regions, decreases as a function of redshift. We observe a systematic decrement not only internally to the VVDS sample, but also when we compare our measurements with the $z = 0$ estimate. This trend is caused by the development of the low- δ tail in the PDF as a function of time on both the $R = 8, 10 h^{-1} \text{Mpc}$ scales and reflects the fact that the probability of having underdense regions is greater at present epoch than it was at $z \sim 1.5$ (where its measured value is $\sim 2\sigma$ lower.)

The amplitudes of the rms and skewness of galaxy overdensities show an evolutionary trend dissimilar from that predicted in first and second order perturbation theory for the gravitational growth of dark matter fluctuations (see Bernardeau et al. 2002 for a review). According to linear perturbation theory the amplitude of the rms of mass fluctuations scales with redshift as in Eq. (18) while second order perturbation theory predicts that, on the scales where the quasi-linear approximation holds, the growth rate of $\langle \delta_g^3 \rangle$ and variance $\langle \delta_g^2 \rangle^2$ are synchronized so that the skewness S_3 of an initially Gaussian fields should remain constant

(Peebles 1980; Juszkiewicz et al. 1993; Bernardeau 1993)¹. Furthermore, in Le Fèvre et al. (2005c) we show that even the general shape of the galaxy PDF deviates from a lognormal distribution, i.e., from the profile in terms of which the mass PDF is generally approximated (see Sect. 5). Therefore, we conclude that the PDF evolution is not caused by gravity alone; the redshift scaling of its global shape and moments effectively indicates the presence of a time evolving bias.

We can deconvolve the purely gravitational signature and investigate properties and characteristics of the biasing between matter and galaxies by comparing the galaxy PDF to the corresponding statistics computed for mass fluctuations. Thus, we now turn to the problem of deriving the PDF of mass fluctuations.

5. The PDF of mass fluctuations in redshift-distorted comoving coordinates

The VVDS survey is providing a rich body of redshift data for mapping the galaxy density field in extended regions of space and over a wide interval of cosmic epochs. On the contrary, the direct determination of the underlying mass density field and its associated PDF is a less straightforward process. Nonetheless we may gain insight into the mass statistics by using simulations and theoretical arguments.

In the standard picture of gravitational instability, the PDF of the primordial cosmological mass density fluctuations is assumed to obey a random Gaussian distribution. Once the density fluctuations reach the non-linear stage, their PDF significantly deviates from the initial Gaussian profile and a variety of phenomenological models have been proposed to describe its shape (e.g., Saslaw 1985; Lahav et al. 1993). In particular, it is well established in CDM models that when structure formation has reached the nonlinear regime, the density contrasts in comoving space $f(\delta)$ follow, to a good approximation, a log-normal distribution (Coles & Jones 1991; Kofman et al. 1994; Taylor & Watts 2000; Kayo et al. 2001),

$$f(\delta) = \frac{(2\pi\omega^2)^{-1/2}}{1+\delta} \exp \left\{ -\frac{[\ln(1+\delta) + \omega^2/2]^2}{2\omega^2} \right\}. \quad (16)$$

This approximation becomes poor in the highly non-linear regime (e.g., Bernardeau & Kofman 1995; Ueda & Yokoyama 1996). The PDF of mass overdensities $f(\delta)$ is characterized by a single parameter (ω) that is related to the variance of the δ -field as

$$\omega^2 = \ln[1 + \langle \delta^2 \rangle]. \quad (17)$$

At high redshifts, the variance σ_R over sufficiently large scales R (those explored in this paper) may be easily derived using the linear theory approximation:

$$\sigma_R(z) = \sigma_R(z=0)D(z) \quad (18)$$

where $D(z)$ is the linear growth rate of density fluctuations normalized to unity at $z = 0$ (Heat 1977; Hamilton 2001).

¹ Note that the observed redshift evolution of the skewness is just the opposite of what is expected also in generic dimensional non-Gaussian models where S_3 is predicted to increase with redshift.

The lognormal approximation formally describes the distribution of matter fluctuations computed in real comoving coordinates. On the contrary, the PDF of galaxies is observationally derived in redshift space. In order to map properly the mass overdensities into galaxy overdensities the mass and galaxy PDFs must be computed in a common reference frame. It has been shown by Sigad et al. (2000) that an optimal strategy to derive galaxy biasing is to compare both mass and galaxy density fields directly in redshift space. Implicit in this approach is the assumption that mass and galaxies are statistically affected in the same way by gravitational perturbations, and thus, that there is no velocity bias in the motion of the two components.

A general model which allows the explicit computation of the statistical distortions caused by peculiar velocities has been proposed by Kaiser (1987). This applies in the linear regime (i.e., on large scales) and in the local Universe where redshift and distances are linearly related. At cosmological distances z , however, the mapping between real comoving coordinates (\mathbf{x}) and redshift comoving coordinates (\mathbf{y}), i.e., the pseudo-comoving coordinates inferred on the basis of the observed redshifts, is less trivial, and we proceed to obtain it in the following.

In an inhomogeneous Universe, galaxies have motions above and beyond their Hubble velocity (e.g., Giovanelli et al. 1998; Marinoni et al. 1998; Branchini et al. 2001). As a consequence, Doppler spectral shifts add to the cosmological signal and the observed redshift (\tilde{z}) is given by

$$\tilde{z} = z + \frac{U(\mathbf{x})}{c}(1+z) \quad (19)$$

where z is the cosmological redshift in a uniform Friedman-Robertson-Walker metric and where $U(\mathbf{x}) = \mathbf{v}(\mathbf{x}) \cdot \hat{\mathbf{r}} = |\mathbf{v}(\mathbf{x})|\mu$ is the radial component of the peculiar velocity (μ is the cosine of the angle between the peculiar velocity vector and the line-of-sight versor $\hat{\mathbf{r}}$).

The redshift comoving distance of a galaxy at the observed redshift \tilde{z} is thus

$$y = \frac{c}{H_0} \int_0^{\tilde{z} + \frac{U}{c}(1+\tilde{z})} \frac{1}{E(\chi)} d\chi, \quad (20)$$

where

$$E(z) = [\Omega_m(1+z)^3 + (1 - \Omega_m - \Omega_\Lambda)(1+z)^2 + \Omega_\Lambda]^{1/2}. \quad (21)$$

At high redshifts ($z \gg \frac{|U|}{c}$), we can write

$$y = \frac{c}{H_0} \left[\int_0^z \frac{1}{E(\chi)} d\chi + \frac{U}{c}(1+z)E(z)^{-1} \right] \quad (22)$$

which, in turns, gives the coordinate transformation from real comoving space \mathbf{x} to the redshift comoving space \mathbf{y}

$$\mathbf{y} = \mathbf{x} \left[1 + p(z) \frac{U(\mathbf{x})}{x} \right]. \quad (23)$$

In this mapping, the cosmological term $p(z)$,

$$p(z) = \frac{1+z}{H_0 E(z)} \quad (24)$$

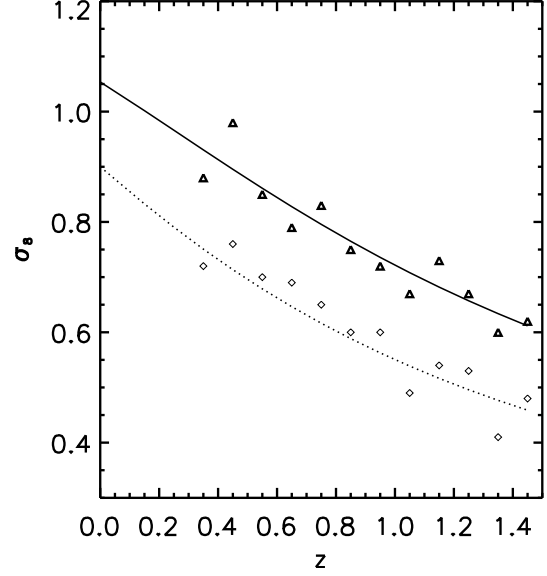


Fig. 8. Redshift scaling of the rms mass fluctuations in sphere of $8 h^{-1}\text{Mpc}$ radius. Diamonds represent σ_8 as computed from the ΛCDM Hubble volume simulation in real comoving space (x -space), while triangles represent the corresponding values recovered in the redshift perturbed comoving coordinates (y -space). The solid line is the analytical prediction for the scaling of σ_8 in the y -space obtained using Eq. (30), while the dotted line represents the x -space evolution predicted in real space. In both cases the power spectrum of perturbations is the same and has been normalized in order to match the simulation specifications ($\sigma^x(z=0) = 0.9$).

is a correcting factor which takes into account the fact that, at high redshifts, distances do not scale linearly with redshift, and, thus, that peculiar velocities cannot be simply added to redshift space positions as in the local Universe.

The galaxy density field in the redshift-distorted space is related to the galaxy density in real space by the Jacobian of the transformation between the two coordinate systems

$$\rho_y(\mathbf{y}) = \rho_x(\mathbf{x}) \left[1 + p(z) \frac{U(\mathbf{x})}{x} \right]^{-2} \left[1 + p(z) \frac{dU(\mathbf{x})}{dx} \right]^{-1}. \quad (25)$$

At sufficiently large distances from the observer, neglecting the survey selection function (i.e., considering a volume-limited redshift survey) and at first order in perturbations we obtain

$$\delta_y(\mathbf{y}) = \delta_x(\mathbf{x}) - p(z) \frac{dU(\mathbf{x})}{dx}. \quad (26)$$

The second term on the right hand side can be evaluated using linear-regime approximations and gravitational instability theory. In comoving coordinates it is given by

$$\frac{dU(\mathbf{x})}{dx} = -\frac{\mu^2 f(z) H(z)}{1+z} \delta_x(\mathbf{x}) \quad (27)$$

where $f = d \ln D / d \ln a$ is the logarithmic derivative of the linear growth rate of density fluctuations with respect to the expansion factor $a(t)$. At redshift z (corresponding to the comoving position \mathbf{x}) a useful approximation is given by:

$$f(z) \sim \Omega_m^{3/5} E(z)^{-6/5} (1+z)^{9/5} \quad (28)$$

(see Martel 1991; Lahav et al. 1991).

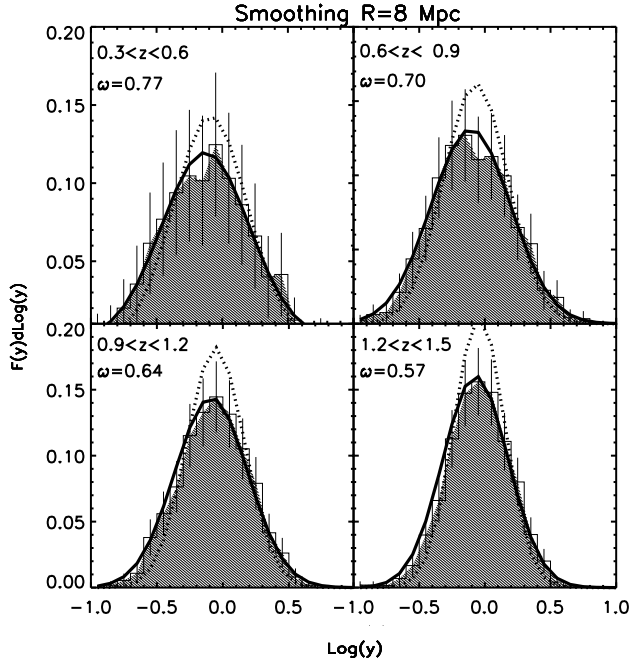


Fig. 9. One-point PDFs of dark matter fluctuations (shaded area) computed using the Hubble volume Λ CDM cosmological simulation in 4 different redshift ranges over a volume which mimics the geometry of the VVDS sample. The mass PDF has been recovered in the redshift comoving space by smoothing the mass-particle distribution with a TH window of size $R = 8 h^{-1}$ Mpc. Note that the plotted histogram actually corresponds to $F(y) = (\ln 10) y f(y)$ because the binning ($d \log y = 0.1$) is done in $\log(y) = \log(1 + \delta)$. The dotted line represents the lognormal approximation derived in the real comoving space using Eq. (16). The solid curve represents the lognormal approximation computed adopting the variance parameter ω (shown in the inset) theoretically inferred using Eq. (30), which models peculiar velocity distortions as a function of redshift.

By combining the previous results we obtain

$$\delta_y(\mathbf{y}) = \delta_x(\mathbf{x})[1 + \mu^2 f(z)]. \quad (29)$$

which reduces to the Kaiser (1987) correction when $z = 0$.

The relation between the azimuthally averaged variances measured in real and redshift comoving space is

$$\sigma^y(z) = \left[1 + \frac{2}{3} f(z) + \frac{1}{5} f^2(z) \right]^{1/2} \sigma^x(z). \quad (30)$$

We have tested the validity of Eq. (30) in the high redshift domain using the Hubble volume N -body simulations carried out by the Virgo consortium (Colberg et al. 2000). This is a large numerical experiment which allows the simulation of mass surveys along the observer past light cone. The simulated mass distribution is computed in a Λ CDM cosmogony with parameters $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$ and $H_0 = 70 \text{ km s}^{-1} \text{ Mpc}^{-1}$. The volume covered by this N -body simulation is large enough that the mass survey extracted along the diagonal of the simulation cubes extends up to the redshift of interest i.e., the redshift covered by the first-epoch VVDS data ($z = 1.5$). In this simulation the mass-particle resolution is $2.2 \times 10^{12} h^{-1} M_\odot$ and the present epoch is defined by a linear rms density fluctuation in a sphere of radius $8 h^{-1}$ Mpc of $\sigma_8 = 0.9$.

The mass density contrasts in the redshift perturbed comoving coordinates $\delta(\mathbf{y}_i, R)$ have been calculated at random positions \mathbf{y}_i in the simulation volume, by smoothing the particle distribution with a spherical top hat window of length $R = 8 h^{-1}$ Mpc. Mass variances in different redshift bins are then derived using Eq. (15). The result is compared to the prediction of Eq. (30) in Fig. 8. Note that, even if it is clear that measurements suffer from cosmic variance due to the relatively small volume sampled at each redshift, the predictions of Eq. (30) are in agreement with the observed scaling of the linear mass variance. The magnitude of the correction with respect to the unperturbed case is also evident; mass fluctuations recovered in redshift space on a $8 h^{-1}$ Mpc scale, in the redshift comoving coordinates, at $z = 0.5(1.5)$ are $\sim 25(35)\%$ larger than in real comoving space (the correction factor is $\sim 17\%$ in the local Universe). Figure 9 shows that this apparent enhancement in the rms fluctuations results in a broadening of the mass PDF recovered in the redshift comoving space. Thus, the effect of peculiar velocities is to shrink overdense regions and to inflate underdense regions, enhancing the probability of having large density fluctuations (both positive and negative).

We finally compare, in various redshift intervals, the accuracy with which the lognormal mass PDF derived in the redshift comoving space (by using Eq. (30) in 17) approximates the PDF directly inferred from the Hubble volume simulation (see Fig. 9). On a scale of $8 h^{-1}$ Mpc, the agreement between the analytical and simulated mass PDFs is satisfactory at all redshifts. This holds true also when the mass PDFs recovered on $R = 5$ and $10 h^{-1}$ Mpc scales are compared.

Thus, with a good degree of confidence, we can use Eq. (30) to predict the PDF of mass fluctuations in redshift distorted comoving coordinates (the same coordinates where the galaxy PDF is observed) and in a generic cosmological background. This allows us to speed up computation time and to frame the results about the biasing function in a generic cosmological model.

6. Measuring galaxy biasing

In this section we describe the method applied to determine the relationship between galaxy and mass overdensities. The galaxy overdensity field δ_g depends in principle on various astrophysical and cosmological parameters such as spatial position (r), underlying matter density fluctuations (δ), scale R with which the density field is reconstructed, cosmological time (z), galaxy colors, local gas temperature, non-local environment, etc.

For the purposes of this study, we will rely on the following simplifying theoretical assumptions:

- i) the efficiency of galaxy formation on a given cosmological scale is sensitive only to the underlying mass distribution. This means that the galaxy fluctuation field is in a reasonably tight one-to-one relationship with the underlying mass fluctuation field, and that the biasing scheme may be formally represented via the relationship $\delta_g = b(z, \delta, R)\delta$. While such an approach represents a non-trivial step forward in understanding the properties of the biasing

function b (if compared, for example, to constant parameterizations of the biasing relation), it is however evident that the biasing function could show, in principle, a more complex functional dependence.

- ii) The current theoretical understanding of how clustering of DM proceeds via gravitational instability in the expanding Universe is well developed, i.e., the PDF of mass fluctuations of the real Universe can be safely derived via analytical models or N -body simulations (see discussion in Sect. 6) In particular, in what follows, we will consider a Λ CDM background mass distribution locally normalized to $\sigma_8(z=0) = 0.9$.
- iii) The redshift distortions affect the densities of galaxies and mass in a similar way, i.e., there is no velocity bias between these two components, and galaxies follow the matter flow.

6.1. The method

As described in Sect. 1, we derive the relationship between galaxy and mass overdensities in redshift space $\delta_g = \delta_g(\delta)$ as the one-to-one transformation which maps the theoretical mass PDF $f(\delta)$ into the observed galaxy PDF $g(\delta_g)$. A similar method to derive the biasing function has been proposed and tested using CDM simulations by Sigad et al. (2000) (see also Szapudi & Pan 2004). This same technique has been recently applied in different contexts by Marinoni & Hudson (2002) to derive the mass-to-light ($M = M(L)$) and the X-ray-to-optical ($L_x = L_x(L)$) functions for a wide mass range of virialized systems, and by Ostriker et al. (2003) to explore the void phenomena in the context of hydrodynamic simulations.

Using Eq. (1)–(3), we obtain the biasing function $b(\delta)$ as the solution of the following differential equation

$$\begin{cases} \delta_g(-1) = -1 \\ b'(\delta)\delta + b(\delta) = f(\delta)g(\delta_g)^{-1} \end{cases} \quad (31)$$

where the prime denotes the derivative with respect to δ , $f(\delta)$ and $g(\delta)$ are the PDF of mass and galaxy fluctuations respectively, and the initial condition has been physically specified by requiring that galaxies cannot form where there is no mass.

With this approach, we loose information on a possible stochasticity characterizing the biasing function. The advantage is that we can provide a measure, on some characteristic scales R , of the *local, non-linear, deterministic* biasing function (Eq. (2)) over the continuous redshift interval $0.4 < z < 1.5$.

We have obtained the biasing function $b(\delta)$ by numerically integrating the differential Eq. (31), i) in different redshift intervals in order to follow the evolution of $b(\delta)$ as a function of cosmic time, and ii) using matter and galaxy PDFs obtained by smoothing the density fields on $R = 5, 8$, and $10 h^{-1}\text{Mpc}$ in order to test the scale dependence of the galaxy biasing function.

The information contained in the non-linear function $b(\delta)$ can be compressed into a single scalar which may be easily compared to the constant values in term of which the biasing relation is usually parameterized (see Eq. (1)). Since, by definition, $\langle b(\delta)\delta \rangle = 0$, the most interesting linear bias estimators are associated to the second order moments of the PDFs, i.e., the variance $\langle \delta_g^2 \rangle$ and the covariance $\langle \delta_g \delta \rangle$. Following the

prescriptions of Dekel & Lahav (1999), we characterize the biasing function as follows:

$$\hat{b} \equiv \frac{\langle b(\delta)\delta^2 \rangle}{\langle \delta^2 \rangle} \quad (32)$$

and

$$b_L^2 \equiv \frac{\langle b^2(\delta)\delta^2 \rangle}{\langle \delta^2 \rangle} \quad (33)$$

where the parameter \hat{b} , measuring the slope of the linear regression of δ_g on δ , is the natural generalization of the linear bias parameter defined in equation 1 and b_L^2 is an “unbiased estimator” of the linear biasing parameter defined as $\xi_g = b^2\xi$, when the bias relation is deterministic, i.e., non-stochastic. The ratio $r = \hat{b}/b_L$ is the relevant measure of nonlinearity in the biasing relation; it is unity for linear biasing, and it is either larger or smaller than unity for nonlinear biasing.

The errors in the measured values of the biasing parameter b_L have been computed using independent mock catalogs which implement all the selection functions of the VVDS. This allows us to incorporate in our error estimates the uncertainties due to cosmic variance.

6.2. Testing the method

Before applying the biasing computation scheme (Eq. (31)) to VVDS data, we have tested that the method can be meaningfully applied, i.e., it is free of systematics when implemented with samples of simulated galaxies which mimic all the observational systematics of our sample.

The procedure consists of computing the biasing function $\delta_s = b(\delta_p)\delta_p$ between the density field δ_s reconstructed using an s-sample (representing the pseudo-survey sample, see Sect. 4.1) and the density fluctuations δ_p of the corresponding p-sample (representing the pseudo-real Universe). We have already determined the range of redshift, density contrasts and smoothing scales where the sample simulating all the VVDS selection functions (s-sample) trace the underlying density of galaxies (p-sample). We thus expect, for consistency, that, in that range, the biasing between the two samples derived by applying our computation scheme (Eq. (31), using the PDFs of the s- and p-samples) is independent of δ_p and equal to $b(\delta_p) = 1$.

Results are presented in Fig. 10 for two different TH smoothing scales. Note that a log-log density plot is used in order to emphasize the behavior of the biasing function in underdense regions. We conclude that on scales $R \geq 8 h^{-1}\text{Mpc}$ the density recovered by a “four-passes” VVDS-like survey is not biased with respect to the underlying distribution on any density scale and in any redshift interval up to $z = 1.5$. As a matter of fact, the linear bias parameter with which information contained in the biasing function can be at first order approximated is $b_L \sim 1$ and the biasing relation does not show any significant deviation from linearity as indicated by the fact that the r parameter is also very close to unity.

If the density field is smoothed on $5 h^{-1}\text{Mpc}$, the effects of the incompleteness in low-density regions (already discussed

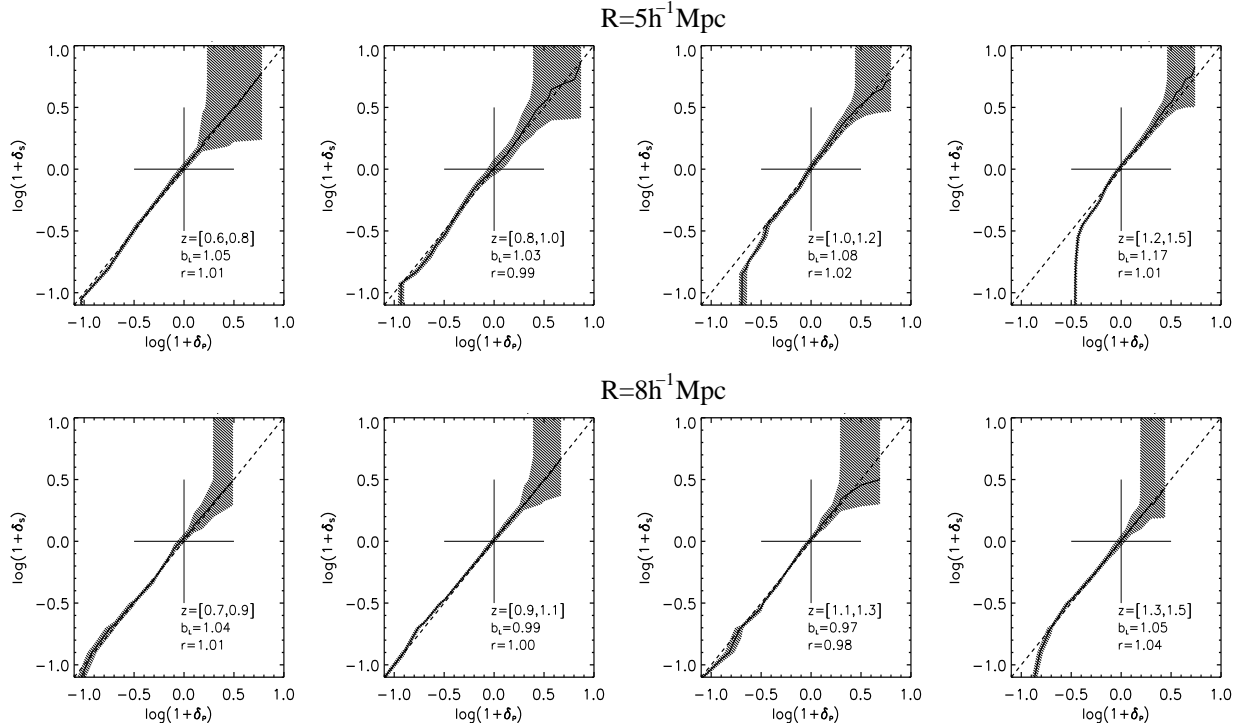


Fig. 10. The simulated biasing function (solid-line) at different cosmic epochs, between the density field traced by the s-sample (GALICS data simulating the VVDS sample, see Pollo et al. (2005) and Sect. 4.1) and the density field traced by the p-sample (GALICS data simulating the real underlying distribution of galaxies). b_L represent the linear bias parameter evaluated from the nonlinear biasing function using the estimator given in Eq. (32). The dashed line is drawn at $b_L = 1$ and represents the no bias case. The central cross is for reference and represents the $\delta_g = \delta = 0$ case. The r parameter measures the deviations from the linearity. The galaxy overdensities are reconstructed using a TH window of sizes $R = 5 h^{-1}\text{Mpc}$ (upper panel) and $R = 8 h^{-1}\text{Mpc}$ (lower panel). The shadowed area represents 1σ errors in the derived biasing function.

in Sect. 4.1, see Figs. 4 and 5) become evident. Underdense regions ($\log(1 + \delta_p) \lesssim -0.5$) in volumes at redshift greater than 1 are poorly sampled with the VVDS survey strategy.

In the same spirit, we have also solved Eq. (31) to determine the biasing relation between the PDF of the Hubble volume mass fluctuations and the lognormal approximation given in Eq. (16). The biasing relation between these two different descriptions of the mass density field is linear and consistent with the no-bias hypothesis between the two representations of the density field on the scales we are interested in ($R \geq 5 h^{-1}\text{Mpc}$ and $\log(1 + \delta) > -1$).

7. The biasing function up to $z \sim 1.5$

7.1. Results

The numerical solutions of Eq. (31) for the $M_B^c = -20 + 5 \log h$ volume-limited VVDS sample are plotted in various redshift slices, in Fig. 11 for the cases $R = 8$ and $10 h^{-1}\text{Mpc}$. Note that a log-log density plot is used in order to emphasize the behavior of the biasing function in underdense regions (note that in these units linear biasing appears as a curved line).

The corresponding parameters b_L and r (also computed for the whole flux-limited sample) are quoted in Table 1, together with our estimates of the second moment of the galaxy PDF σ_R and of the skewness parameter S_3 . Both these statistics have been computed as described in Sect. 5. Also note that the values of σ_8 measured for the flux-limited sample are consistent with

the values independently derived in Sect. 3.4 on the basis of the results of the analysis of the clustering properties of VVDS galaxies (Paper III).

An empirical fit of the biasing function is obtained by using a formula similar to the one proposed by Dekel & Lahav (1999)

$$\delta_g(\delta) = \begin{cases} (1 + a_0)(1 + \delta)^{a_1} - 1 & \delta \leq 0 \\ a_0 + a_2\delta + a_3\delta^2 & \delta > 0 \end{cases} \quad (34)$$

which best describes the behavior of biasing in underdense regions ($\delta < 0$), either the second order Taylor expansion of the density contrast of dark matter (Fry & Gaztañaga 1993)

$$\delta_g = \sum_{k=0}^2 \frac{b_k}{k!} \delta^k \quad (35)$$

which allows an easier comparison of our results with other studies. The best fitting parameters of these non-linear approximations are quoted in Table 2.

The dependence of the shape of the biasing function on galaxy luminosity is plotted in Fig. 12. Results are shown at the median depth of the VVDS sample (in the redshift bin $0.7 < z < 0.9$) where faint objects ($M_B < -17.7 + 5 \log h$) are still sampled.

In Fig. 13 we show the redshift evolution of the linear biasing parameter b_L computed over the redshift interval $0.4 < z < 1.5$ for both the flux and volume limited samples. We can conclude that biasing is not changing with cosmic time for

Table 1. Bias measurements from the VVDS first epoch data.

R $h^{-1}\text{Mpc}$	Redshift range	\mathcal{M}_B^c	N_{gal}	b_L	r	σ_R	S_3
5	$0.4 < z < 0.7$	No	1583	0.87 ± 0.15	0.95	0.94 ± 0.15	1.2 ± 0.3
	$0.7 < z < 0.9$		1044	0.95 ± 0.15	0.96	0.98 ± 0.15	1.4 ± 0.3
	$0.9 < z < 1.1$		759	0.97 ± 0.13	0.97	0.93 ± 0.13	1.1 ± 0.3
5	$0.4 < z < 0.7$	-18.7	610	1.06 ± 0.17	0.97	1.18 ± 0.17	1.6 ± 0.3
	$0.7 < z < 0.9$		726	1.03 ± 0.15	0.97	1.05 ± 0.15	1.6 ± 0.3
	$0.9 < z < 1.1$		751	1.00 ± 0.14	0.97	0.97 ± 0.14	1.5 ± 0.3
5	$0.4 < z < 0.7$	-20	160	1.10 ± 0.18	0.97	1.28 ± 0.18	1.4 ± 0.4
	$0.7 < z < 0.9$		229	1.12 ± 0.17	0.99	1.18 ± 0.17	1.1 ± 0.3
	$0.9 < z < 1.1$		289	1.18 ± 0.15	0.97	1.17 ± 0.15	1.4 ± 0.4
8	$0.7 < z < 0.9$	No	1263	0.92 ± 0.20	0.97	0.67 ± 0.20	1.6 ± 0.4
	$0.9 < z < 1.1$		864	1.03 ± 0.16	0.96	0.74 ± 0.16	1.6 ± 0.3
	$1.1 < z < 1.3$		440	1.21 ± 0.12	0.96	0.82 ± 0.12	1.3 ± 0.3
	$1.3 < z < 1.5$		234	1.51 ± 0.10	0.95	0.96 ± 0.11	1.0 ± 0.3
8	$0.7 < z < 0.9$	-18.7	879	0.98 ± 0.21	0.98	0.75 ± 0.21	1.6 ± 0.4
	$0.9 < z < 1.1$		813	1.01 ± 0.16	0.97	0.72 ± 0.16	1.8 ± 0.3
8	$0.7 < z < 0.9$	-20	279	1.17 ± 0.23	0.98	0.89 ± 0.23	1.7 ± 0.4
	$0.9 < z < 1.1$		327	1.30 ± 0.17	0.99	0.94 ± 0.17	1.8 ± 0.3
	$1.1 < z < 1.3$		251	1.33 ± 0.12	0.96	0.88 ± 0.12	1.2 ± 0.3
	$1.3 < z < 1.5$		169	1.56 ± 0.11	0.95	0.99 ± 0.11	1.1 ± 0.3
10	$0.7 < z < 0.9$	No	1425	1.03 ± 0.22	0.91	0.66 ± 0.22	1.6 ± 0.4
	$0.9 < z < 1.1$		955	1.05 ± 0.18	0.97	0.64 ± 0.18	1.7 ± 0.4
	$1.1 < z < 1.3$		480	1.17 ± 0.13	0.90	0.68 ± 0.13	1.3 ± 0.3
	$1.3 < z < 1.5$		250	1.55 ± 0.11	0.93	0.84 ± 0.13	1.2 ± 0.3
10	$0.7 < z < 0.9$	-18.7	991	1.03 ± 0.25	0.95	0.69 ± 0.25	1.5 ± 0.4
	$0.9 < z < 1.1$		900	1.03 ± 0.18	0.96	0.63 ± 0.18	1.7 ± 0.3
10	$0.7 < z < 0.9$	-20	316	1.14 ± 0.25	0.92	0.75 ± 0.25	1.6 ± 0.4
	$0.9 < z < 1.1$		360	1.26 ± 0.20	0.97	0.78 ± 0.20	1.8 ± 0.4
	$1.1 < z < 1.3$		266	1.36 ± 0.14	0.91	0.78 ± 0.14	1.3 ± 0.3
	$1.3 < z < 1.5$		175	1.54 ± 0.13	0.93	0.84 ± 0.13	1.3 ± 0.3

Table 2. Best fitting parameters of the non linear biasing models given in Eqs. (34) and (35). Errors do not include cosmic variance.

R $h^{-1}\text{Mpc}$	Redshift range	\mathcal{M}_B^c	a_0	a_1	a_2	a_3	b_0	b_1	b_2
8	$0.7 < z < 0.9$	No	0.36 ± 0.03	1.36 ± 0.11	0.60 ± 0.12	-0.01 ± 0.02	0.10 ± 0.06	0.88 ± 0.10	-0.12 ± 0.08
	$0.9 < z < 1.1$		0.35 ± 0.03	1.36 ± 0.40	1.08 ± 0.10	-0.08 ± 0.02	0.23 ± 0.06	1.18 ± 0.09	-0.20 ± 0.06
	$1.1 < z < 1.3$		0.39 ± 0.04	1.64 ± 0.13	1.14 ± 0.11	-0.08 ± 0.02	0.24 ± 0.07	1.26 ± 0.10	-0.22 ± 0.06
	$1.3 < z < 1.5$		0.54 ± 0.05	2.34 ± 0.30	1.46 ± 0.14	-0.11 ± 0.02	0.37 ± 0.12	1.57 ± 0.15	-0.24 ± 0.08
8	$0.7 < z < 0.9$	-20	0.25 ± 0.05	1.67 ± 0.20	1.25 ± 0.16	-0.09 ± 0.03	0.18 ± 0.10	1.30 ± 0.15	-0.20 ± 0.08
	$0.9 < z < 1.1$		0.40 ± 0.06	2.12 ± 0.33	1.20 ± 0.16	-0.05 ± 0.02	0.23 ± 0.17	1.29 ± 0.16	-0.12 ± 0.08
	$1.1 < z < 1.3$		0.45 ± 0.04	1.86 ± 0.13	1.29 ± 0.13	-0.11 ± 0.02	0.31 ± 0.08	1.40 ± 0.13	-0.26 ± 0.06
	$1.3 < z < 1.5$		0.48 ± 0.05	2.51 ± 0.30	1.46 ± 0.16	-0.10 ± 0.02	0.33 ± 0.14	1.55 ± 0.15	-0.22 ± 0.08
10	$0.7 < z < 0.9$	No	0.25 ± 0.03	0.85 ± 0.10	1.44 ± 0.21	-0.23 ± 0.06	0.17 ± 0.07	1.20 ± 0.15	-0.30 ± 0.12
	$0.9 < z < 1.1$		0.18 ± 0.02	1.20 ± 0.08	1.23 ± 0.13	-0.18 ± 0.04	0.20 ± 0.04	1.19 ± 0.09	-0.34 ± 0.08
	$1.1 < z < 1.3$		0.39 ± 0.04	1.30 ± 0.11	1.04 ± 0.13	-0.08 ± 0.03	0.26 ± 0.06	1.18 ± 0.10	-0.22 ± 0.08
	$1.3 < z < 1.5$		0.54 ± 0.04	2.09 ± 0.13	1.41 ± 0.15	-0.14 ± 0.03	0.39 ± 0.08	1.55 ± 0.12	-0.32 ± 0.08
10	$0.7 < z < 0.9$	-20	0.21 ± 0.03	1.33 ± 0.10	1.45 ± 0.16	-0.22 ± 0.04	0.18 ± 0.06	1.26 ± 0.11	-0.28 ± 0.10
	$0.9 < z < 1.1$		0.18 ± 0.04	1.51 ± 0.15	1.38 ± 0.16	-0.13 ± 0.03	0.20 ± 0.08	1.36 ± 0.14	-0.26 ± 0.08
	$1.1 < z < 1.3$		0.38 ± 0.04	1.37 ± 0.17	1.50 ± 0.15	-0.13 ± 0.03	0.37 ± 0.10	1.50 ± 0.14	-0.26 ± 0.08
	$1.3 < z < 1.5$		0.22 ± 0.05	1.76 ± 0.20	1.83 ± 0.16	-0.18 ± 0.03	0.33 ± 0.12	1.73 ± 0.17	-0.34 ± 0.08

$z < 0.8$, while there is a more pronounced evolution of biasing in the redshift interval $[0.8, 1.5]$. In particular, the difference between the value of b_L at redshift $z \sim 1.5$ and $z \sim 0$ for a population of galaxies with luminosity $\mathcal{M}_B < -20 + 5 \log h$ is $\Delta b_L \sim 0.5 \pm 0.14$, thus significant at a confidence level greater than 3σ .

In Fig. 14 we show the dependence of the linear biasing parameter on galaxy luminosity. Intrinsically brighter galaxies are more strongly biased than less luminous ones at every redshift and the dependence of biasing on luminosity at $z \sim 0.8$ is in good agreement with what is observed in the local Universe (Norberg et al. 2001).

Given the difference in the rest-frame colors of elliptical and irregular galaxies and the fact that the observed I band corresponds to bluer rest-frame bands at higher redshift, the relative fraction of early- and late-type galaxies in our I band limited survey will change as a function of redshift.

Specifically, the observed difference in the B -band luminosity function of early- and late-types (Zucca et al. 2005), implies that the VVDS survey selects preferentially late-type galaxies at higher redshift. It is known that at $z = 0$ late-type galaxies cluster less strongly than early-types (e.g., Giovanelli et al. 1986; Guzzo et al. 1997; Giuricin et al. 2001; Madgwick et al. 2002; Zehavi et al. 2002), and, thus, we might observe a variation of the amplitude of density fluctuations at high redshifts just because the morphological composition of our sample changes.

In order to disentangle the spurious morphological contribution to the observed evolution of the global biasing function we have splitted our sample according to rest frame colors, selecting a red ($(B - I)_0 > 0.95$; 849 galaxies in the 4-passes region with $z > 0.7$) and a blue subsample of galaxies ($(B - I)_0 < 0.68$; 1891 galaxies with $z > 0.7$). These color cuts roughly correspond to selecting, respectively, morphological types $\leq II$ and IV according to the classification scheme devised by Zucca et al. 2005 for the VVDS sample.

Clearly, this subsample selection does not correspond to the ideal case of a redshift survey sampling galaxies according to their rest-frame colors; however, useful information about differences in clustering between red and blue populations can still be inferred.

Note that the hypothesis on which the technique of comparing mass and galaxy density distributions is based (Sect. 6) can be straightforwardly generalized to compute the biasing between the density distributions of different galaxy types. In particular, we assume that the large scale velocities of late and early types are not dissimilar relative to each other (as it is effectively observed at $z = 0$ e.g., Dekel 1994; Marinoni et al. 1998) i.e., the two velocity fields are noisy versions of the same underlying field.

Results about the color dependence of biasing are summarized in Table 3 and graphically presented in Fig. 15. The red sample is systematically a more biased tracer of mass than the blue one in every redshift interval investigated (i.e., $b^r > b^b$), but the relative biasing between the two populations is nearly constant ($b^r/b^b \sim 1.4 \pm 0.1$)

7.2. Analysis and discussion

7.2.1. Biasing for the global galaxy population

Here we examine and interpret the results derived in the previous section. We begin by discussing the general shape of the non-linear biasing function, for the global population, in different density regions. Our results can be summarized as follows:

i) in underdense regions ($1 + \delta < 1$) the local slope of the biasing function $b(\delta)$ is always larger than unity even when the global slope is $b_L < 1$ (see for example Fig. 12). The fact that galaxies in low-mass density regions are always positively biased with respect to the mass distribution (i.e., locally $b > 1$)

is possibly physically caused by the fact that galaxies do not form in very low-density mass regions, i.e., below some finite mass underdensity the galaxy formation efficiency drops to zero. Using the biasing relation given in Eq. (34) the characteristic mass density threshold δ_c below which very few galaxies form ($\delta_g \leq -0.9$), can be approximated as

$$\log(1 + \delta_c) \sim -\frac{1 + \log(1 + a_0)}{a_1}. \quad (36)$$

There is evidence that this mass-density threshold, characterizing regions avoided by galaxies, increases as a function of redshift (see Fig. 11) and luminosity (see Fig. 12). If we consider $R = 8 h^{-1}\text{Mpc}$ and the redshift bin $0.7 < z < 0.9$ we see that while faint galaxies seem to be present even where the mass density contrast is very low (left panel of Fig. 12, $\log(1 + \delta_c) = -0.96 \pm 0.09$),

brighter galaxies do not seem to form in deep mass underdensities (right panel of Fig. 12, $\log(1 + \delta_c) = -0.73 \pm 0.11$). Therefore low-density regions are preferentially inhabited by low luminosity galaxies.

Moreover the mass-density threshold below which the formation of bright galaxies ($M_B < -20 + 5 \log h$) seems to be inhibited increases, irrespective of the scale investigated (see Fig. 11) as a function of redshift. On a scale $R = 8 h^{-1}\text{Mpc}$, the threshold shifts from $\log(1 + \delta_c) = -0.73 \pm 0.10$ at $z = 0.8$ to $\log(1 + \delta_c) \sim -0.55 \pm 0.07$ at $z = 1.4$. This suggests that galaxies of a given luminosity were tracing systematically higher mass overdensities in the early Universe, i.e., as time progresses, galaxy formation begins to take place also in lower density peaks.

ii) Even in regions where the mass density distribution is close to its mean value ($1 + \delta \sim 1$) bright galaxies are not unbiased tracer of the mass-overdensity field (Fig. 11). This can also be seen by setting $\delta = 0$ in Eq. (34) and noting that $\delta_g(\delta = 0) = a_0 > 0$ for both analyzed samples (flux- and volume-limited) in all redshift ranges (see Table 2). This result is at variance with what is expected within the simple linear biasing picture, where, by construction, $\delta_g(\delta = 0) = 0$.

iii) In higher matter-density environments ($1 + \delta > 1$) galaxies were progressively more biased mass tracers in the past, i.e., the local slope $b(\delta)$ systematically increases with redshift on every scale investigated (Fig. 11). There is some indication that, at the upper tail of the mass density distribution, galaxies are anti-biased with respect to mass on all scales (i.e., the local slope is $b(\delta) < 1$ for $\delta \gg 1$). Antibiasing in overdense regimes is a feature actually observed in simulations (e.g., Sigad et al. 2000; Somerville et al. 2001) and expected in theoretical models (e.g., Taruya & Sato 2000). Physically this could be due to the merging of galaxies which reduces the number density of visible objects in high density regions or because galaxy formation is inhibited in regions where the gas is too hot to collapse and form stars.

iv) In general the linear approximation offers a poor description of the richness of details encoded in the biasing function. As a matter of fact the linear biasing function (dotted line in Figs. 11 and 12) poorly describes, in many cases, the observed scaling of the biasing relation (solid line). At the comoving scales of $R = 5, 8$ and $10 h^{-1}\text{Mpc}$, non-linearities in the

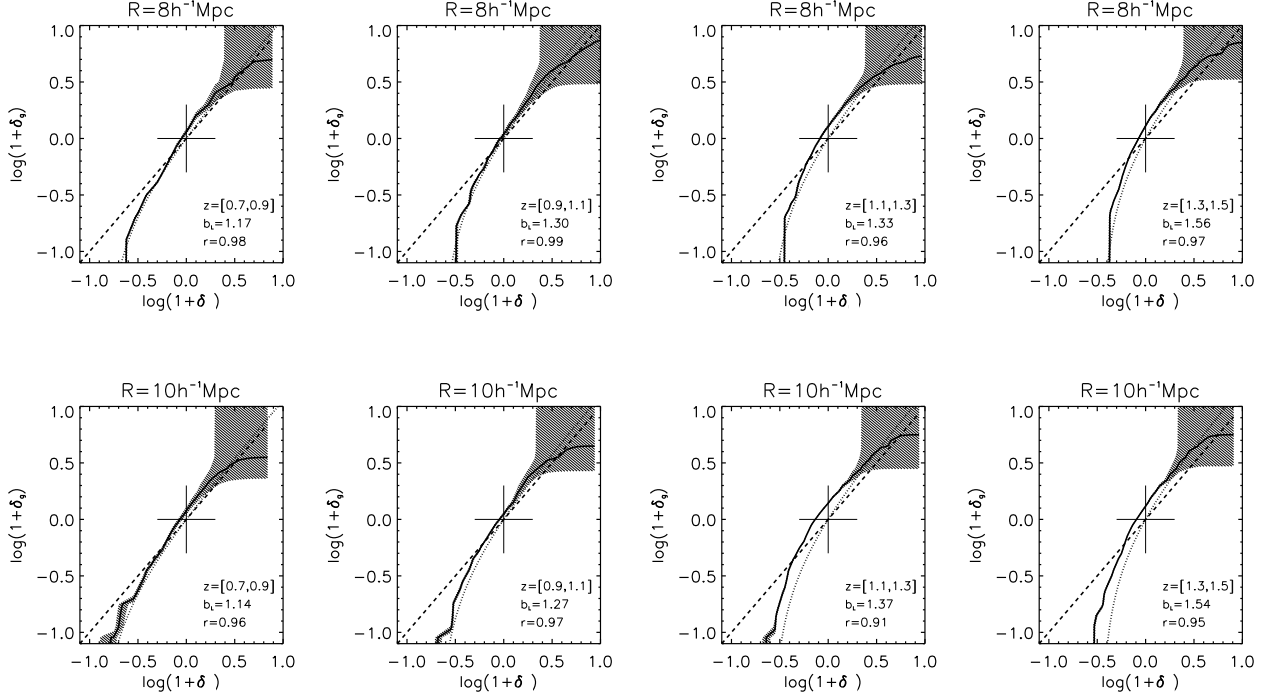


Fig. 11. The observed biasing function (solid-line) recovered for the density field smoothed on scales $R = 8 h^{-1}\text{Mpc}$ (upper panel) and $10 h^{-1}\text{Mpc}$ (lower panel) and for different redshift bins (from left to right) in the volume-limited VVDS sample ($M_B^c = -20 + 5 \log h$). The dotted line represents the linear biasing model $\delta_g = b_L \delta$ while the no-bias case ($b_L = 1$) is shown with a dashed line. The central cross is for reference and represents the $\delta_g = \delta = 0$ case. The shaded area represents 1σ errors in the derived biasing function. Errors take into account the noise in the observed galaxy PDF ($g(\delta_g)$), but do not include uncertainties due to cosmic variance.

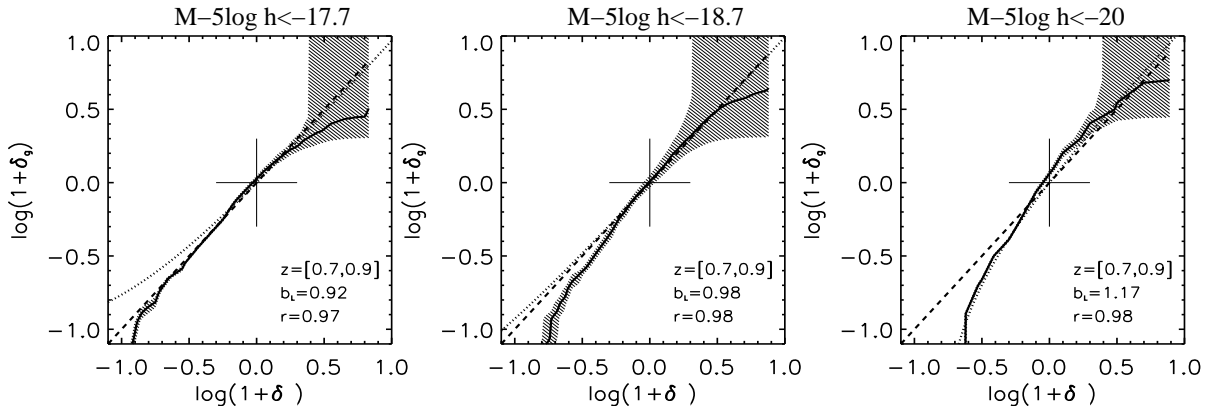


Fig. 12. The biasing function (solid-line) on scales $R = 8 h^{-1}\text{Mpc}$ and in the redshift interval $0.7 < z < 0.9$ computed for different luminosity classes. Symbols are as in Fig. 11.

biasing relation are typically $(1-r) \lesssim 10\%$ in the redshift range investigated. We find that the ratio b_2/b_1 between the quadratic and linear term of the series approximation given in Eq. (35) is nearly constant in the redshift range $0.7 < z < 1.5$ and does not depend on luminosity (i.e. it is nearly the same for the flux- and volume-limited subsamples) or smoothing scale. We find that, on average, $b_2/b_1 \sim -0.15 \pm 0.04$ for $R = 8 h^{-1}\text{Mpc}$ and $b_2/b_1 \sim -0.19 \pm 0.04$ for $R = 10 h^{-1}\text{Mpc}$.

To facilitate comparison with other studies, which generally focus on the linear representation of biasing, we now discuss the properties of the linear approximation of our biasing function. The general characteristics of the linear parameter b_L can be summarized as follows:

v) by inspecting Table 1, we do not find any significant evidence that the global value of the linear biasing

parameter b_L depends on the smoothing scale. Any possible systematic variation, if present, is smaller than the amplitude of our errorbars (~ 0.15). This scale independence in the biasing relation extends into the high redshift regimes similar conclusions obtained in the local Universe by the 2dFGRS on scales $> 5 h^{-1}\text{Mpc}$ (Verde et al. 2002). Moreover our results may be interpreted as a supporting evidence for theoretical arguments suggesting that bias is expected to be scale-independent on scales larger than a few $h^{-1}\text{Mpc}$ (e.g., Mann et al. 1998; Weinberg et al. 2004).

Since we find no evidence of scale-dependent bias, and since with different R scales we are probing different redshift regimes, in Fig. 13 we have averaged the linear biasing parameters measured on 5, 8, and $10 h^{-1}\text{Mpc}$ scales (values quoted in Table 1) in order to follow, in a continuous way, the

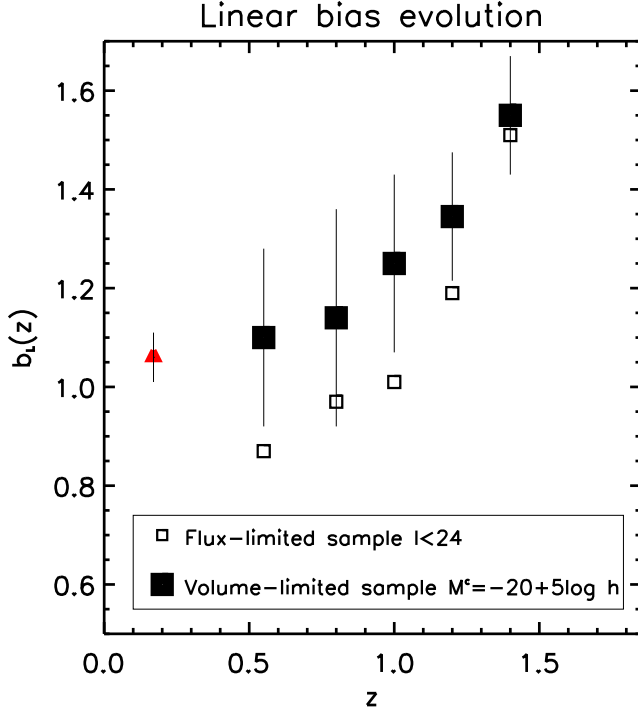


Fig. 13. The redshift evolution of the linear biasing parameter b_L for the volume-limited ($M_B^c < -20 + 5 \log h$) subsample (filled squares) is compared to the evolution of the biasing parameter for the whole flux-limited VVDS-02h sample (empty squares). Since there is no significant evidence of scale dependence in the biasing relation, we have averaged the biasing parameters measured on 5, 8, and 10 h^{-1} Mpc scales in order to cover the full redshift baseline $0.4 < z < 1.5$. For clarity, only the errorbars corresponding to the volume-limited sample are shown. The triangle represents the $z \sim 0$ bias inferred for 2dFGRS galaxies having median $L/L^* \sim 2$ (i.e., the median luminosity of the volume-limited VVDS sample) as explained in the text.

redshift evolution of the linear galaxy biasing over the larger redshift baseline $0.4 < z < 1.5$. Figure 13 shows that b_L for galaxies brighter than $M_B = -20 + 5 \log h$ changes from 1.10 ± 0.18 at $z \sim 0.55$ to 1.55 ± 0.12 at $z \sim 1.4$.

An even steeper variation is observed for the biasing of the flux-limited sample, indicating that biasing depends on galaxy luminosity. Figure 13 shows that the ratio between the amplitude of galaxy fluctuations and the underlying mass fluctuations declines with cosmic time. This scaling is effectively predicted within the framework of the peaks-biasing theoretical model (Kaiser 1984). At early times, galaxies are expected to form at the highest peaks of the density field since one needs a dense enough clump of baryons in order to start forming stars. Such high- σ peaks are highly biased tracers of the underlying mass density field. According to this picture, as time progresses and the density field evolves, galaxy formation moves to lower- σ peaks, nonlinear peaks become less rare events and thus galaxies become less biased tracers of the mass density field. Additional “debiasing” mechanisms may contribute to the observed scaling shown in Fig. 13. It is likely that the densest regions stop forming new galaxies because their gas becomes too hot, cannot cool efficiently, and thus cannot collapse and form stars (Blanton et al. 1999). As galaxy formation moves

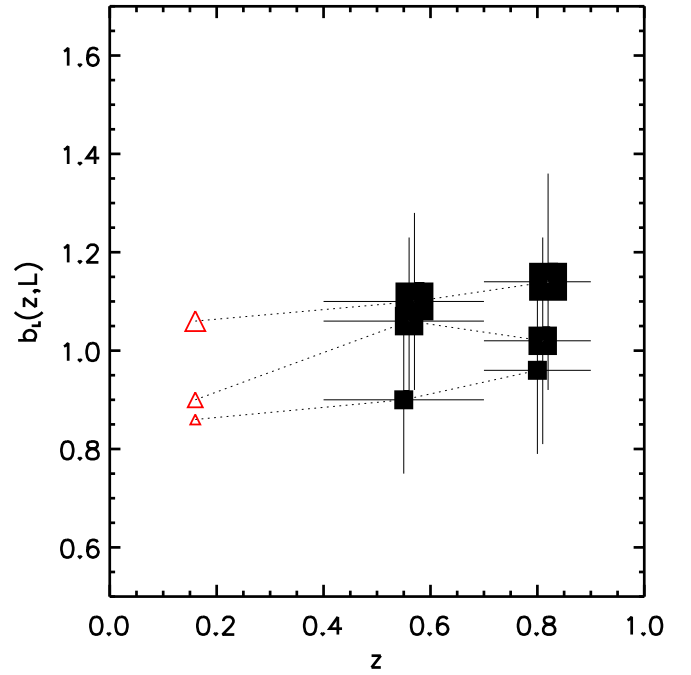


Fig. 14. Comparison between the galaxy linear bias parameter measured in the redshift interval $0.4 < z < 0.9$ for 3 different luminosity classes (squares) and the corresponding local estimates provided by the 2dFGRS (triangles). Points with increasing sizes correspond to three different volume-limited VVDS subsamples, i.e. $M_B - 5 \log h < -17.7, < -18.7$ and < -20 , respectively. For clarity, squares with increasing size have been progressively displaced rightward to avoid crowding. The $z \sim 0$ measurements have been interpolated by using the formula describing the luminosity dependence of the 2dFGRS bias parameter (Norberg et al. 2001), the bias parameter for the 2dFGRS L^* sample (i.e., $b_* = 0.92$ (Verde et al. 2002)) and the median luminosity of the three VVDS subsamples ($L/L^* = 0.52, 0.82, 2.0$ respectively).

out of the hottest (and rarest) regions of the Universe, the biasing decreases. Finally, we also note that in order to derive the biasing function we have assumed that there is no difference in the velocity field of the luminous and matter components. After galaxies form, they are subject to the same gravitational forces as the dark matter, and thus they tend to trace the dark matter distribution more closely with time as shown by Dekel & Rees (1987); Fry (1996); Tegmark & Peebles (1998).

vi) In Fig. 13 we also show, for comparison, the value of the 2dFGRS linear biasing parameter inferred at $z = 0.17$ (the effective depth of the survey) as the ratio between the σ_8 value measured by Croton et al. 2004 (in redshift-distorted space; see their Figs. 3 and 4) for a sample of objects with $-21 < M_B - 5 \log h < -20$ (which actually brackets the median luminosity of our volume limited sample $\sim 2L^*$), and the rms of mass fluctuations (in redshift-distorted space) in a Λ CDM background (see Sect. 6). This value ($\sim 1.07 \pm 0.06$) is in excellent agreement with what one would independently obtain by combining the linear bias parameter measured by Verde et al. (2002) for the whole 2dFGRS (1.04 ± 0.11) with the bias scaling law recipe of Norberg et al. (2001), i.e., $b(z = 0.17, L = 2L^*) = 1.07 \pm 0.13$.

We can conclude that the time dependence of biasing is marginal ($db/b \sim 7 \pm 25\%$) for $z < 0.8$ while it is substantial ($db/b \sim 33 \pm 18\%$) in the redshift interval $[0.8-1.5]$. The observed time evolution of bias is well described by the simple scaling relationship $b_L = 1 + (0.03 \pm 0.01)(1+z)^{3.3 \pm 0.6}$ in the interval $0 < z < 1.5$.

Assuming a linear biasing scheme, one may note that this result was already implicit in Fig. 7 of Sect. 4. The rms fluctuations of the mass density field on a $8 h^{-1}\text{Mpc}$ scale decrease monotonically with redshift by a factor of $\sim 22\%$ and $\sim 23\%$ in the redshift intervals $[0.17-0.8]$ and $[0.8-1.4]$, respectively; thus, a nearly constant bias is predicted in the redshift range $z = [0.17-0.8]$ because the rms fluctuations of the galaxy density field are also decreasing by a factor $\sim 16\%$ in this same interval. Since, instead, σ_8 of galaxies is marginally increasing in the range $z = 0.8-1.4$ ($d\sigma/\sigma \sim 10\%$, see Table 1), over this redshift baseline the biasing evolves rapidly.

vii) Bright galaxies are more biased mass tracers than the general population (see Fig. 12). This result confirms and extends into the high redshift domain the luminosity dependence of biasing which is observed in local samples of galaxies (e.g., Benoist et al. 1996; Giuricin et al 2001; Norberg et al. 2001; Zehavi et al. 2002). Specifically, in Fig. 14 we show the dependence of galaxy biasing from luminosity measured in the redshift interval $0.4 < z < 0.9$ using three different volume-limited VVDS subsamples (i.e., $M_B - 5 \log h < -17.7, < -18.7$ and < -20 respectively) and compare their linear biasing parameters with those observed locally for a sample of objects having the same median luminosities of the VVDS subsamples (i.e., $L/L^* = 0.52, 0.82, 2.0$ respectively). The local estimates have been computed on the basis of the scaling relationship $b/b^* = 0.85 + 0.15 L/L^*$ derived by Norberg et al. (2001) using the 2dFGRS sample, assuming the b^* value given by Verde et al. (2002). As shown above for the volume-limited sample, no significant evolution is seen up to $z \sim 0.8$ also when the dependence of bias from luminosity is analyzed.

Finally, we note that, as already discussed in Sect. 5, galaxies with the same luminosity at different redshifts may actually correspond to different populations. Since, as we have shown, biasing increases with luminosity also at high redshift, and since the measured value of M^* for our sample at redshift $z = 0.4(1.5)$ (Paper II) is fainter(brighter) than the cut-off magnitude $M_B^c = -20 + 5 \log h$, we can infer that $b_L(z)$ for a population of objects selected, at any given redshift, in a narrow luminosity range around $M^*(z)$ should increase with redshift even more than what we have measured for our volume-limited sample (see Fig. 13). A more detailed analysis of the biasing for $M^*(z)$ galaxies will be presented in the future, when a larger VVDS data sample will be available.

7.2.2. Biasing as a function of galaxy color

Results summarized in Table 3 and presented in Fig. 15 show that, on scales $R = 8 h^{-1}\text{Mpc}$, the red sample is a more biased tracer of mass than the blue one in every redshift interval. Similarly to what we have found for the global population, there is some indication of a systematic increase as a function

Table 3. The biasing parameters for red and blue VVDS subsamples on a scale $R = 8 h^{-1}\text{Mpc}$.

Redshift range	Volume limited	b_L^{rel}	b_L^R	b_L^B
$0.7 < z < 0.9$	No	1.3 ± 0.2	1.3 ± 0.5	1.0 ± 0.5
$0.9 < z < 1.1$	No	1.3 ± 0.2	1.4 ± 0.5	1.1 ± 0.4
$1.1 < z < 1.3$	No	1.4 ± 0.2	1.6 ± 0.5	1.2 ± 0.4
$1.3 < z < 1.5$	No	1.4 ± 0.3	2.3 ± 0.4	1.6 ± 0.4
$0.7 < z < 0.9$	-20	1.4 ± 0.3	1.5 ± 0.6	1.1 ± 0.6
$0.9 < z < 1.1$	-20	1.3 ± 0.3	1.6 ± 0.6	1.2 ± 0.5
$1.1 < z < 1.3$	-20	1.5 ± 0.2	2.0 ± 0.5	1.3 ± 0.4
$1.3 < z < 1.5$	-20	1.3 ± 0.3	2.1 ± 0.4	1.6 ± 0.4

of redshift of the biasing of bright red and blue objects even if, because of the large errorbars, this trend is not statistically significant.

We can compare our results to the biasing measured for extremely red objects (EROS), i.e., objects with extremely red colors ($(R - K)_{\text{vega}} > 5$). Using the results of the correlation analysis of Firth et al. (2002), we obtain, for their $(I - H)_{\text{vega}} > 3$, $H_{\text{vega}} < 20.5$ sample (which has a median blue luminosity $M_B = -20.3 + 5 \log h$), $b_L^{\text{EROS}}(z \sim 1.2) \sim 2.3 \pm 0.6$. Considering the results of Daddi et al. (2001), who analyzed a sample of EROS with $(R - K)_{\text{vega}} > 5$ (which roughly corresponds to $(I - H)_{\text{vega}} > 3$), $K_{\text{vega}} < 19.2$ sample, we conclude that $b_L^{\text{EROS}}(z \sim 1.2) \sim 4 \pm 1$. These values for the galaxy biasing are respectively ~ 0.5 and 1.8σ higher than that measured for our sample of bright ($M_B < -20 + 5 \log h$) but moderately red galaxies ($b_L^r(z \sim 1.2) = 2 \pm 0.5$). One may interpret this results as an indication for the reddest objects being more strongly biased than moderately red galaxies of similar luminosity. Anyway, given the large errorbars, the evidence that, at $z \sim 1.2$, the biasing properties of these two differently selected populations are different is not statistically significant. As a matter of fact, the values quoted above are also consistent with an alternative hypothesis, i.e., the strength of the EROS fluctuations with respect to the mass fluctuations is not exceptional when compared to the density fluctuations observed in a sample of high redshift, moderately red galaxies of similar luminosity.

The specific values of the biasing parameter at each cosmic epoch are affected by large errors due to the sparseness of our volume-limited subsamples, and to the presence of cosmic variance. One way to bypass uncertainties due to cosmic variance consists in computing the relative biasing function $b^{\text{rel}}(\delta) = b^r(\delta)/b^b(\delta)$ between the red and blue subsamples. As the subsamples are drawn from the same volume, this ratio should be minimally affected by the finiteness of the volume probed by the first epoch VVDS data.

Results about the relative biasing between galaxy of different colors are graphically shown in the lower panel of Fig. 15, while estimates of the corresponding b_L are quoted in Table 3.

We do not observe any trend in the relative biasing between red and blue volume-limited subsamples in the redshift range $0.7 < z < 1.5$. Moreover, our best estimate $b_L^{\text{rel}} \sim 1.4 \pm 0.1$ is in excellent agreement with what is found for nearly the same color-selected populations both locally (Willmer et al. (1998)

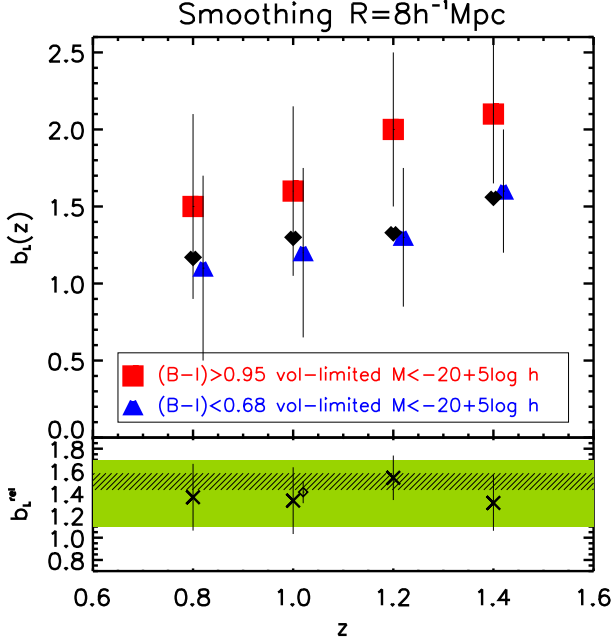


Fig. 15. *Upper panel:* redshift evolution of the galaxy bias on a scale $R = 8 h^{-1}\text{Mpc}$ for the red (squares) and blue (triangles) galaxies in the volume limited samples. For clarity, the triangles have been slightly displaced rightward to avoid crowding. Black diamonds represent the global bias for galaxies brighter than $M_B^c = -20 + 5 \log h$. *Lower panel:* the relative bias between the red and blue population ($b^{\text{rel}} = b^r/b^b$) is shown as a function of redshift. The filled and shaded areas represent the 1σ confidence region of the $z \sim 0$ value for the relative bias derived by Wilmer et al. (1998) and Wild et al. (2005) respectively. The diamond represents the relative bias measured by Coil et al. (2004) in the redshift interval (0.7–1.35).

found that, on a scale $R = 8 h^{-1}\text{Mpc}$, $b_L^{\text{rel}} \equiv b((B-R)_{0,\text{vega}} > 1.3)/b((B-R)_{0,\text{vega}} < 1.3) = 1.4 \pm 0.3$, while Wild et al. (2005) using the 2dFGRS found on the same scale $b_L^{\text{rel}} \equiv b((B-R)_{0,\text{vega}} > 1.07)/b((B-R)_{0,\text{vega}} < 1.07) = 1.5 \pm 0.07$ and at $z \sim 1$ (Coil et al. (2004) found, on a scale $R = 8 h^{-1}\text{Mpc}$, that $b_L^{\text{rel}} \equiv b((B-R)_0 > 0.7)/b((B-R)_0 < 0.7) = 1.41 \pm 0.10$). Thus, VVDS results suggest that there is no-redshift dependence for the relative biasing between red and blue objects up to $z \sim 1.5$. Possible systematics could conspire to produce the observed results; the linear approximation may not always captures, in an accurate way, all the information contained in the biasing function, and more importantly, a purely magnitude limited survey samples the red and blue populations at high redshift with a different efficiency (see discussion in Sect. 7.1).

In principle, the relative bias could be further studied as a function of scale. For example, locally, there is evidence of scale dependence in the relative bias with the bias decreasing as scale increases (Wilmer et al. 1998, Madgwick et al. 2003, Wild et al. 2005). However, the sample currently available is not sufficiently large to obtain proper statistics on this effect, although this should be measurable from the final data set.

Finally, we note that no differences in the value of b_L^{rel} are seen by comparing volume-limited subsamples with the flux-limited one in different redshift intervals (see Table 3).

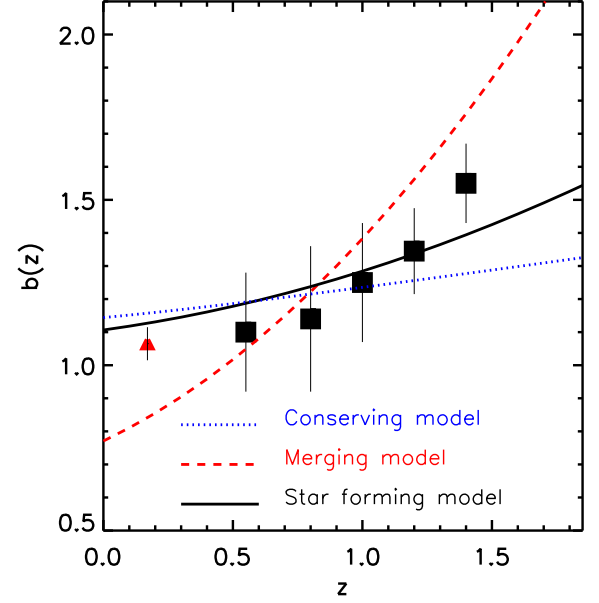


Fig. 16. The redshift evolution of the linear biasing parameter b_L for the volume-limited ($M_B^c = -20 + 5 \log h$) sample (see Fig. 13) is compared to various theoretical models of biasing evolution. The dotted line indicates the conserving model normalized at $b_f(z = 1.4) = 1.28$, the solid and dashed lines represent the star forming and merging models with the mass thresholds set at 3.2×10^{11} and $2.4 \times 10^{12} h^{-1} M_\odot$ respectively.

Thus, we can deduce that in each redshift bins $b^R(M_B < -20)/b^R \sim b^B(M_B < -20)/b^B$. In other terms the biasing between the most luminous objects of a particular color and the global population of objects of the same type appears to be independent of galaxy colors (see Table 3).

8. Comparison with theoretical predictions

In this section we compare our results about the biasing of the $M_B^c = -20 + 5 \log h$ volume-limited, global galaxy sample, with predictions of different theoretical models.

Since we have found that the distribution of galaxy and mass fluctuations are different and the bias was systematically stronger in the past, we can immediately exclude the scenario in which galaxies trace the mass at all cosmic epochs. We thus consider more complex theoretical descriptions of the biasing functions, in particular three different pictures based on orthogonal ideas of how evolution proceeds: the *conserving*, the *merging*, and the *star forming* biasing models (see e.g., Moscardini et al. 1998).

In the first model the number of galaxies is conserved as a function of time (Dekel & Rees 1987; Fry 1996). This model does not assume anything about the distribution and mass of dark matter halos or their connection with galaxies. In this scheme one assumes that galaxies are biased at birth and then they follow the flow of matter without merging, in other terms they behave as test particles dragged around by the surrounding density fluctuations. Because the acceleration on galaxies is the same as that on the dark matter, the gravitational evolution after formation will tend to bring the bias closer to unity, as described by Fry (1996) and Tegmark & Peebles (1998).

Table 4. Best fitting parameters and the corresponding χ^2 values for various biasing models.

Model	Best fitting parameters	$\chi^2/\text{d.o.f.}$
Conserving	$b_f(z = 1.4) = 1.28 \pm 0.03$	2
Merging	$M = 2.4 \times 10^{12} h^{-1} M_\odot$	5.5
Star forming	$M = (3.2 \pm 3) \times 10^{10} h^{-1} M_\odot$	0.7

The evolution of the bias is given by (e.g., Tegmark & Peebles 1998)

$$b(z) = 1 + (b_f - 1) \frac{D(z_f)}{D(z)} \quad (37)$$

where b_f is the bias at the formation time z_f .

An alternative picture for the bias evolution, which explicitly takes into account galaxy merging, has been proposed by Mo & White (1996) who gave analytical prescriptions for computing the bias of halos using the Press & Schechter formalism.

If we explicitly assume that galaxies can be identified with dark matter halos, an approximate expression for the biasing of all halos of mass $> M$ existing at redshift z (but which collapsed at redshift greater than the observation redshift, see discussion in Matarrese et al. 1997) is given by

$$b(M, z) = 1 + \frac{1}{\delta_c} \left(\frac{\delta_c^2}{\sigma^2(M, z)} - 1 \right) \quad (38)$$

where $\delta_c \sim 1.69$ is the linear overdensity of a sphere which collapses in an Einstein-de Sitter Universe and $\sigma(M, z)$ is the linear rms fluctuations on scales corresponding to mass M at the redshift of observation.

The third model is also framed within the peaks-biasing formalism. It assumes that the distribution of galaxies with luminosity $> L$ is well traced by halos with mass $> M$, and predicts the biasing of objects that just collapsed at the redshift of observation (e.g., Blanton et al. 2000). In this *star forming* model,

$$b(M, z) = 1 + \frac{\delta_c}{\sigma^2(M, z)} \quad (39)$$

represents the biasing of galaxies that formed in a narrow time interval around redshift z (i.e., galaxies which experienced recent star formation at redshift z).

Clearly the above models, are based on a set of theoretical ingredients which represent a crude approximation of the complex multiplicity of physical phenomena entering the cosmic recipe of galaxy biasing. In this context, our goal is to investigate the robustness of the simplifying assumptions on which theoretical models are based, and explore the validity or limits of their underlying physical motivations.

Theoretical predictions are compared to observations (VVDS data plus the local normalization derived from 2dFGRS data) in Fig. 16. The best fitting parameters for each model are evaluated using a χ^2 statistics and are quoted, together with the corresponding minimum χ^2 value of the fit, in Table 4.

The best fitting galaxy conserving model is obtained when the bias at birth is $b_f(z_f = 1.4) = 1.28 \pm 0.03$ and the corresponding normalized χ^2 -value is $\chi_N^2 = 2$. As shown in Fig. 16 the redshift evolution predicted by this model is much weaker

than suggested by data. Thus, the gravitational debiasing is a physical mechanism that alone may not fully explain the observed redshift evolution of the biasing, in the sense that it significantly underpredicts the rate of evolution.

The redshift evolution is more pronounced in the *merging* model (specifically, in Fig. 16, we show the bias evolution of galaxies hosted in halos having mass $M \gtrsim 2.4 \times 10^{12} h^{-1} M_\odot$). While this model successfully describes the redshift dependence of the biasing of halos (Mo & White 1996; Somerville et al. 2001) it poorly accounts for the redshift evolution of the bias of galaxies (with $M_B \leq -20 + 5 \log h$) between $z = 0$ and $z = 1.5$ which is slower than predicted ($\chi_N^2 = 5.5$ for the best fitting model). Thus, although merging is an important mechanism for describing the evolution of matter clustering, our result implies that merging processes affect galaxies in a less dramatic way than halos. Since in the Press & Schechter formalism halos are required to merge instantaneously in bigger units at the redshift of observation, our result would imply, also, that the merger time-scales of galaxies is different from that of halos. Moreover, selecting galaxies with a fixed luminosity threshold may not correspond, over such a wide z range as that investigated here, to selecting halos above a given fixed mass threshold. In this sense our result would be suggestive of evolution in the mass-to-light ratio as a function of time.

In Fig. 16 we also show the expected redshift evolution for the *star forming* model for halos of $M \gtrsim 3.2 \times 10^{10} h^{-1} M_\odot$. In this case, the agreement between model and observations is better ($\chi_N^2 = 0.7$). Clearly this does not mean that we are analyzing a sample of objects that just collapsed and formed stars at the time they were observed; as a matter of fact the model cannot capture all the physical processes shaping the biasing relation. Moreover, the low value fitted for the mass threshold is somewhat unrealistic for the bright objects we are considering. Notwithstanding, Blanton et al. (2000) already noted that the prediction of this biasing model is not much different from the biasing evolution expected for the general population of galaxies in a hydrodynamical simulation of the large scale structure.

Our analysis seems to suggest the apparent need of more complex biasing models that better approximate the observed biasing evolution. Understanding our results completely, however, will require more discriminatory power in the data, and, thus, a larger VVDS sample.

9. Summary and conclusions

Deep surveys of the Universe provide the basic ingredients needed to compute the probability distribution function of galaxy fluctuations and to constrain its evolution with cosmic time. The evolution of the galaxy PDF may shed light onto the general assumption that structures grows via gravitational collapse of density fluctuations that are small at early times. When this statistic is combined with analytical CDM predictions for the PDF of mass, useful insights into the biasing function relating mass and galaxy distributions can be obtained.

In this paper, we have explored the potentiality of this approach by analyzing the first-epoch data of the VVDS survey. This is the largest, purely flux-limited sample of

spectroscopically measured galaxies, currently available in a continuously connected volume and with a robust sampling up to redshifts $z \sim 1.5$. The VVDS is probing the high redshift domain at $I \leq 24$ in the VVDS-02h-4 field with the same sampling rate of pioneer surveys of the local Universe such as the CFA (at $z \sim 0$) and, more recently, the 2dFGRS (at $z \sim 0.1$).

Particular attention has been paid to assess the completeness of the VVDS sample and to test the statistical reliability of the PDF of VVDS galaxy fluctuations. In particular:

- a) by applying the VVDS observational selection functions to GALICS semi-analytical galaxy simulations we have explored the region of the parameter space where the PDF of VVDS-like densities traces in a statistically unbiased way the parent underlying PDF of the real distribution of galaxy overdensities.
- b) we have reconstructed the VVDS galaxy density field on different scales $R = [5, 8, 10] h^{-1}\text{Mpc}$ by correcting the density estimator for various VVDS selection functions. The final density map for the flux-limited sample has been Wiener filtered in order to minimize the shot-noise contribution.

By studying the PDF of galaxies in the high redshift Universe we have found that the peak of the galaxy PDF systematically shifts to lower density contrasts as a function of redshift and that the probability of observing underdense regions is greater at $z \sim 0.7$ than it was at $z \sim 1.5$. Both these effects provide strong supporting evidence for the standard assumption that the large scale structure is the result of the gravitational growth of small primordial density fluctuations in an expanding universe.

First, within the paradigm of gravitational instability, the assembling process of the large-scale structures is thought to be regulated by the interplay of two competing effects: the tendency of local self-gravity to make overdense regions collapse and the opposite tendency of global cosmological expansion to move them apart. A key signature of gravitational evolution of density fluctuations in an expanding Universe is that underdense regions, experiencing the cosmological matter outflow, occupy a larger volume fraction at present epoch than in the early Universe. Secondly, both these effects, the peak shift and the development of a low density tail, could indicate the existence of a time-evolving biasing between matter and galaxies, since galaxy biasing, systematically increasing with redshift, offers a natural mechanism to re-map the galaxy PDF into progressively higher intervals of density contrasts.

This last interpretation is confirmed by our measurements of the evolution properties of the second and third moments of the galaxy PDF. We find that i) the rms amplitude of the fluctuations of bright VVDS galaxies is with good approximation constant over the full redshift baseline investigated. Specifically we have shown that, in redshift space, σ_8 for galaxies brighter than $M_B^c = -20 + 5 \log h$ has a mean value of 0.94 ± 0.07 in the redshift interval $0.7 < z < 1.5$; ii) the third moment of the PDF, i.e., the skewness, increases with cosmic time. Its value at $z \sim 1.5$ is nearly 2σ lower than measured locally by the 2dFGRS. Both these results, when compared to predictions of linear and second order perturbation theory, unambiguously indicate that galaxy biasing is an increasing function of redshift.

Exploiting the sensitivity of the galaxy PDF to the specific form of the mass-galaxy mapping, we have derived the redshift-, density-, and scale-dependent biasing function $b(z, \delta, R)$ between galaxy and matter fluctuations in a Λ CDM Universe, by analyzing the Jacobian transformation between their respective PDFs. Particular attention has been paid to devise an optimal strategy so that the comparison of the PDFs of mass and galaxies can be carried out in an objective and accurate way. Specifically, we have corrected the lognormal approximation, which describes the mass density PDF, in order to take into account redshift distortions induced by galaxy peculiar velocities at early cosmic epochs where the mapping between redshifts and comoving positions is not linear. In this way, theoretical predictions can be directly compared to observational quantities derived in redshift space.

Without a priori parameterizing the form of the biasing function, we have shown its general non trivial shape, and studied its evolution as a function of cosmic epoch. Our main results about biasing in the high redshift Universe can be summarized as follows:

- i) we detect non-linear effects in the biasing relation. The ratio between the quadratic and linear term of the biasing expansion (cf. Eq. (35)) is different from zero at a confidence level greater than 3σ in all the redshift bins and for all the smoothing scales probed. This result confirms a general prediction of CDM-based hierarchical models of galaxy formation (e.g., Sigad et al. 2000; Somerville et al. 2001). Such non-linear distortions of the biasing function are not observed locally in the 2dFGRS sample, although indirect evidence of a non-linear bias at $z \sim 0$ exists (Benoist et al. 1999, Baugh et al. 2004).
- ii) The biasing function rises sharply in underdense regions (the local slope is $b(\delta) > 1$) indicating that below some finite mass density threshold the formation efficiency of galaxies brighter than $M_B < -20 + 5 \log h$ drops to zero. This threshold shifts towards higher values of the mass density field as the luminosity or the redshift of the galaxy population increases.
- iii) We do not observe the imprints of scale-dependency in the biasing function a behavior in agreement with results derived from more local surveys at $z \sim 0$ (Verde et al. 2002).
- iv) By representing the biasing function in linear approximation, we have found that the linear biasing parameter b_L evolves with cosmic time: it appears that we live in a special epoch in which the galaxy distribution traces the underlying mass distribution on large scales ($b_L \sim 1$), while, in the past, the two fields were progressively dissimilar and the relative biasing systematically higher. The difference between the value of b_L at redshift $z \sim 1.5$ and $z \sim 0$ for a population of galaxies with luminosity $M_B < -20 + 5 \log h$ is significant at a confidence level greater than 3σ ($\Delta b_L \sim 0.5 \pm 0.14$). In this interval, the essential characteristics of the time evolution of the linear bias are well described in terms of the phenomenological relationship $b_L = 1 + (0.03 \pm 0.01)(1 + z)^{3.3+0.6}$.
- v) Over the redshift baseline investigated, the rate of biasing evolution is a function of redshift: $z \sim 0.8$ is the

characteristic redshift which marks the transition from a “minimum-evolution” late epoch to an early period where the biasing evolution for a population of $M_B < -20 + 5 \log h$ galaxies is substantial ($\sim 33 \pm 18\%$ between redshift 0.8 and 1.5).

- vi) Brighter galaxies are more strongly biased than less luminous ones at every redshift and the dependence of biasing on luminosity at $z \sim 0.8$ is in good agreement with what is observed in the local Universe.
- vii) By comparing our results to predictions of theoretical models for the biasing evolution, we have shown that the galaxy *conserving* model (Fry 1996) and halo *merging* (Mo & White 1996) model offer a poor description of our data. This result could suggest that the gravitational debiasing and the hierarchical merging of halos may not be the only physical mechanisms driving the evolution of galaxy biasing across cosmic epochs. At variance with these results, the *star forming* model (Blanton et al. 2000) seems to describe better the observed redshift evolution of the linear biasing factor.
- viii) After splitting the first-epoch redshift catalog into red and blue volume-limited subsamples, we have found that the red sample is systematically a more biased tracer of mass than the blue one in every redshift interval investigated, but the relative biasing between the two populations is nearly constant in the redshift range $0.7 < z < 1.5$ ($b^r/b^b \sim 1.4 \pm 0.1$), and comparable with local estimates. Moreover, we have found that the bright red subsample is biased with respect to the general red population in the same way as the bright sample of blue objects is biased with respect to the global blue population thus indicating that biasing as a function of luminosity might be, at first order, independent of color.
- ix) Because the VVDS and various EROS samples are not yet large enough, the bias of our sample of bright and moderately red objects at $z \sim 1$ is not statistically dissimilar from that expected for EROS of similar luminosity, even if the EROS biasing appears to be systematically larger.

One key aspect of this paper is the measure of evolution in the distribution properties of galaxy overdensities from a continuous volume sampled with the same selection function over a wide redshift baseline. As our volume sampled is still limited, errors in the analysis presented in this paper are dominated by cosmic variance. The technique presented here will be applied to a larger sample as the VVDS observational program progresses.

Acknowledgements. We would like to acknowledge useful discussions with A. Dekel, R. Giovanelli and L. Moscardini. We also thank S. Andreon and J. Afonso for their useful comments on the paper. This research has been developed within the framework of the VVDS consortium and it has been partially supported by the CNRS-INSU and its Programme National de Cosmologie (France), and by the Italian Ministry (MIUR) grants COFIN2000 (MM02037133) and COFIN2003 (num.2003020150). CM also acknowledges financial support from the Region PACA. The VLT-VIMOS observations have been carried out on guaranteed time (GTO) allocated by the European Southern Observatory (ESO) to the VIRMOS consortium, under a

contractual agreement between the Centre National de la Recherche Scientifique of France, heading a consortium of French and Italian institutes, and ESO, to design, manufacture and test the VIMOS instrument. We thank the GALICS group for privileged access to their semi-analytic simulations. The mass simulations used in this paper were carried out by the Virgo Supercomputing Consortium using computers based at the Computing Centre of the Max-Planck Society in Garching and at the Edinburgh parallel Computing Centre.

References

- Adelberger, K. L., Steidel, C. C., Giavalisco, M., et al. 1998, *ApJ*, 505, 18
- Arnouts, S., Cristiani, S., Moscardini, L., et al. 1999, *MNRAS*, 310, 540
- Bagla, J. S. 1998, *MNRAS*, 299, 417
- Basilakos, S., & Plionis, M. 2001, *MNRAS*, 316, 779
- Bardeen, J., Bond, J. R., Kaiser, N., & Szalay, A. 1986, *ApJ*, 304, 15
- Baugh, C. M., Croton, D. J., Gastañaga, E., et al. 2004, *MNRAS*, 351, L44
- Benoist, C., Maurogordato, S., da Costa, L. N., Cappi, A., & Schaeffer, R. 1996, *ApJ*, 472, 452
- Benoist, C., Cappi, A., da Costa, L. N., et al. 1999, *ApJ*, 514, 563
- Bernardeau, F., Colombi S., Gaztanaga E., & Scoccimarro, R. 2002, *Phys. Rep.* 367, 1
- Bernardeau, F. 1994, *ApJ*, 433, 1
- Bernardeau, F., & Kofman, L. 1995, *ApJ*, 443, 479
- Blaizot, J., Wadadekar, Y., Guiderdoni, B., et al. 2003, [[arXiv:astro-ph/0309305](https://arxiv.org/abs/astro-ph/0309305)]
- Blanton, M., Cen, R., Ostriker, J. P., & Strauss, M. A. 1999, *ApJ*, 522, 590
- Blanton, M., Cen, R., Ostriker, J. P., Strauss, M. A., & Tegmark, M., 2000, *ApJ*, 531, 1
- Bottini et al. (the VVDS team), 2005, *PASP*, in press, [[astro-ph/0409252](https://arxiv.org/abs/astro-ph/0409252)]
- Branchini, E., Freudling, W., Da Costa, L. N., et al. 2001, *MNRAS*, 326, 1191
- Carlberg, R. G., Yee, H. K. C., Morris, S. L., et al. 2000, *ApJ*, 542, 57
- Catelan, P., Matarrese, S., & Porciani, C. 1998, *ApJ*, 502, L1
- Cen, R., & Ostriker, J. P. 1992, *ApJ*, 399, L113
- Coil, A., Davis, M., Madgwick, D. S., et al. 2004, *ApJ*, 609, 525
- Colberg, J. M., White, S. D. M., Yoshida, N., et al. (The Virgo Consortium), 2000, *MNRAS*, 319, 209
- Coles, P., & Jones, B. 1991, *MNRAS*, 248, 1
- Coles, P. 1993, *MNRAS*, 262, 1065
- Colless, M. M., Dalton, G., Maddox, S., et al. 2001, *MNRAS*, 328, 1039
- Croton, D. J. 2004, *MNRAS*, 352, 1232
- Daddi, E., Broadhurst, T., Zamorani, G., et al. 2001, *A&A*, 376, 825
- Dalal, N., & Kochanek, C. S. 2002, *ApJ*, 572, 25
- Davis, M., & Geller, M. J. 1976, *ApJ*, 254, 437
- Davis, M., & Huchra, J. 1981, *ApJ*, 254, 437
- Davis, M., & Djorgovski, S. 1985, *ApJ*, 299, 15
- Davis, M., Efstathiou, G., Frenk, C., & White, S. D. M. 1985, *ApJ*, 292, 371
- Davis, M., Faber, S. M., Newman, J. et al. 2003, *Proc. SPIE*, 4834, 161
- Dekel, A. 1994, *ARA&A*, 32, 371
- Dekel, A., & Lahav, O. 1999, *ApJ*, 520, 24
- Dekel, A., & Rees, M. J. 1987, *Nature*, 326, 455
- Dressler, A. 1980, *ApJ*, 236, 351
- Evrard, A. E., Summers, F. J., & Davis, M. 1994, *ApJ*, 422, 11

- Firth, A. E., Somerville, R. S., McMahon, R. G. et al. 2002, *MNRAS*, 332, 617
- Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., & Huchra, J. P. 1993, *ApJ*, 402, 42
- Foucaud, S., McCracken, H. J., Le Fèvre, O., et al. 2003, *A&A*, 409, 835
- Fry, J. N. 1985, *ApJ*, 289, 10
- Fry, J. N., & Gaztañaga, E. 1993, *ApJ*, 413, 447
- Fry, J. N. 1996, *ApJ*, 461, L65
- Giavalisco, M., Steidel, C. C., Adelberger, K. L., et al. 1998, *ApJ*, 503, 543
- Giovanelli, R., Haynes, M., & Chincarini G. 1986, *ApJ*, 300, 77
- Giovanelli, R., Haynes, M. P., Freudling, W., et al. 1998, *ApJ*, 505, 91
- Giuricin, G., Samurovic, S., Girardi, M., Mezzetti, M., & Marinoni, C. 2001, *ApJ*, 554, 857
- Guzzo, L., Strauss, M. A., Fisher, K. B., Giovanelli, R., & Haynes, M. P. 1997, *ApJ*, 489, 37
- Hamilton, A. J. S. 2001, *MNRAS*, 322, 419
- Hamilton, A. J. S. 1988, *ApJ*, 331, L59
- Heat, D. J. 1977, *MNRAS*, 179, 351
- Hatton, S., Devriendt, J. E. G., Ninin, S., et al. 2003, *MNRAS*, 373, 75
- Hudson, M. 1993, *MNRAS*, 265, 43
- Hui, L., Gaztañaga, E. 1999, *ApJ*, 519, 622
- Ilbert, O., et al. (the VVDS team) 2005, *A&A*, 439, 863 (Paper II)
- Juszkiewicz, R., Bouchet, F. R., & Colombi, S. 1993, *ApJ*, 412, L9
- Kaiser, N. 1984, *ApJ*, 284, L9
- Kaiser, N. 1987, *MNRAS*, 227, 1
- Kaiser, N., & Peacock, J. A. 1991, *ApJ*, 379, 482
- Kayo, I., Taruya, A., & Suto, Y. 2001, *ApJ*, 561, 22
- Klypin, A., Kravtsov, A. V., Valenzuela, O., & Prada, F. 1999, *ApJ*, 522, 82
- Kofman, L., Bertshinger, E., Gelb, J. M., Nusser, A., & Dekel, A. 1994, *ApJ*, 420, 44
- Kravtsov, A. V., & Klypin A. A. 1999, *ApJ*, 520, 437
- Lahav, O., Itoh, M., Inagaki, S., & Suto, Y. 1993, *ApJ*, 402, 387
- Lahav, O., Fisher, K. B., Hoffman, Y., Scharf, C. A., & Zaroubi, S. 1994, *ApJ*, 423, L93
- Lahav, O., Bridle, S. L., Percival, W. J. et al. 2002, *MNRAS*, 333, 961
- Lahav, O., Lilje, P. B., Primack, J. R., Rees, M. J. et al. 1991, *MNRAS*, 251, 128
- Le Fèvre, O., Hudon, D., Lilly, S. J., et al. 1996, *ApJ*, 461, 534
- Le Fèvre, O., Vettolani, G., Maccagni, D., et al. (the VVDS team) 2003, *The Messenger*, 111, 18
- Le Fèvre, O., Vettolani, G., Paltani, S. et al. (the VVDS team) 2004, *A&A*, 428, 1043
- Le Fèvre, O., et al. (the VVDS team) 2005a *A&A*, 439, 845 (Paper I)
- Le Fèvre, O., et al. (the VVDS team) 2005b, *A&A*, 439, 877 (Paper III)
- Le Fèvre, O., et al. (the VVDS team) 2005c, *Nature*, in press
- Madgwick, D. S., Lahav, O., Baldry, I. K. et al. 2002, *MNRAS*, 333, 133
- Magliocchetti, M., Bagla, J. S., Maddox, S., & Lahav, O. 2000, *MNRAS*, 314, 546
- Mann, R. G., Peacock, J. A., & Heavens, A. F. 1998, *MNRAS*, 293, 209
- Marinoni, C., Monaco, P., Giuricin, G., & Costantini, B. 1998, *ApJ*, 505, 484
- Marinoni, C., & Hudson, M. J. 2002, *ApJ*, 569, 101
- Martel, H. 1991, *ApJ*, 377, 7
- Matarrese, S., Coles, P., Lucchin, F., & Moscardini, L. 1997, *MNRAS*, 286, 115
- Maurogordato, S., & Lachièze-Rey, M. 1987, *ApJ*, 320, 13
- Mc Cracken, H., Radovich, M., Bertin, E., et al. 2003, *A&A*, 410, 17
- Mo, H., & White, S. D. M. 1996, *MNRAS*, 282, 347
- Moore, B. 1999, *ApJ*, 524, L19
- Moscardini, L., Coles, P., Lucchin, F., & Matarrese, S. 1998, *MNRAS*, 299, 95
- Navarro, J. F., Frenk C. S., & White S. D. M. 1997, *ApJ*, 490, 493
- Norberg, P., Baugh, C. M., Hawkins, E., et al. 2001, *MNRAS*, 328, 64
- Oke, J. B., & Gunn, J. E. 1983, *ApJ*, 266, 713
- Ostriker, J. P., Nagamine, K., Cen, R., & Fukugita, M. 2003, *ApJ*, 597, 1
- Peebles, P. J. E. 1980, *The Large Scale Structure of the Universe* (Princeton: Princeton Univ Press)
- Pollo, A., et al. (the VVDS team) 2005, *A&A*, 439, 887
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 1992, *Numerical Recipes*, (Cambridge: University Press)
- Rybicki, G. B., & Press, W. H. 1992, *ApJ*, 398, 169
- Saslaw, W. C., *Gravitational Physics of Stellar and Galactic Systems* (Cambridge: Cambridge Univ. Press)
- Sigad, Y., Branchini, E., & Dekel, A. 2000, *ApJ*, 540, 62
- Small, T. A., Ma, C., Sargent, W. L. W., & Hamilton, D. 1999, *ApJ*, 524, 31
- Somerville, R. S., Lemson, G., Sigad, Y., et al. 2001, *MNRAS*, 320, 289.
- Steidel, C. C., Adelberger, K. L., Dickinson, M. E., et al. 1998, *ApJ*, 492, 428
- Strauss, M. A., & Willick, J. A. 1995, *Phys. Rev.* 261, 271
- Szapudi, I., & Pan, J. 2004, *ApJ*, 602, 26
- Taruya, A., & Suto, Y. 2000, *ApJ*, 542, 559
- Taylor, A. N., & Watts, P. I. R. 2000, *MNRAS*, 314, 92
- Tegmark, M., & Peebles, P. J. E. 1998, *ApJ*, 500, L79
- Ueda, H., & Yokoyama, J. 1996, *MNRAS*, 280, 754
- Verde, L., Heavens, A. F., Percival, W. J., et al. 2002, *MNRAS*, 335, 432
- Weinberg, D. H., Davé, R., Katz, N., & Hernquist, L. 2004, *ApJ*, 601, 1
- White, S. D. M., & Rees, M. J. 1978, *MNRAS*, 183, 341
- White, S. D. M., Tully, B. R., & Davis, M. 1988, *ApJ*, 333, L45
- Wild, V., Peacock, J. A., Lahav, O., et al. 2005, *MNRAS*, 356, 247
- Willmer, C. N. A., da Costa, L. N., & Pellegrini, P. S. 1998, *AJ*, 115, 869
- Zaroubi, S., Hoffman, Y., Fisher, K. B., & Lahav, O. 1995 *ApJ*, 449, 446
- Zehavi, I., Blanton, M. R., Frieman, J. A. et al. 2002, *ApJ*, 571, 172
- Zucca, E. et al. (the VVDS team), 2005, *A&A*, submitted, [arXiv:astro-ph/0506393]

Appendix A

Here we describe the application of the Wiener filtering technique to deconvolve the noise signature from the VVDS density map. We de-noise data in Fourier space, noting, however, that an equivalent filtering can be directly applied in real space (e.g., Rybicki & Press 1992; Zaroubi et al. 1995).

Let us assume that the observed smoothed density field $\delta_O(\mathbf{x})$, and the true underlying density field $\delta_T(\mathbf{x})$, smoothed on the same scale, are related via

$$\delta_O(\mathbf{x}) = \delta_T(\mathbf{x}) + \epsilon(\mathbf{x}), \quad (40)$$

where $\epsilon(\mathbf{x})$ is the local contribution from shot noise (see Eq. (8)). The Wiener filtered density field, in Fourier space, is

$$\tilde{\delta}_F(\mathbf{k}) = \mathcal{F}(\mathbf{k})\tilde{\delta}_O(\mathbf{k}), \quad (41)$$

$$\mathcal{F}(\mathbf{k}) = \frac{\langle \tilde{\delta}_T^2(\mathbf{k}) \rangle}{\langle \tilde{\delta}_T^2(\mathbf{k}) \rangle + (2\pi)^3 P_\epsilon(\mathbf{k})} \quad (42)$$

where brackets denote statistical averages and where $P_\epsilon(\mathbf{k}) = (2\pi)^{-3} \langle |\tilde{\epsilon}^2(\mathbf{k})| \rangle$ is the power spectrum of the noise. Assuming ergodic conditions for the noise, we can derive its power spectrum as $P_\epsilon(\mathbf{k}) = (2\pi)^{-3} |\tilde{\epsilon}(\mathbf{k})|^2$.

The Power spectrum of the underlying theoretical density distribution of galaxies, smoothed with the window F and taking into account the VVDS geometrical constraints, can be derived from Eq. (5). Specifically, the theoretical overdensity field smoothed on a certain scale R , which is sampled by an idealized survey with no selection functions, is

$$\delta_g(\mathbf{r}, R) \equiv \frac{1}{\bar{\rho}} \sum_p \delta^D(\mathbf{r} - \mathbf{r}_p) * F\left(\frac{|\mathbf{r} - \mathbf{r}_p|}{R}\right) - 1. \quad (43)$$

If we assume that this density field is periodic on same volume V , having, for example, the same geometry of the VVDS survey, its Fourier transform is

$$\tilde{\delta}_T(\mathbf{k}) = \frac{1}{\bar{\rho}} \sum_i \frac{n_i}{V} e^{i\mathbf{k} \cdot \mathbf{r}_i} \tilde{F}_k - \frac{1}{V} \int e^{i\mathbf{k} \cdot \mathbf{r}} d^3\mathbf{r} \quad (44)$$

where the mean theoretical galaxy density $\bar{\rho}$ may be estimated averaging over sufficiently large volumes the VVDS data corrected for selection functions and sampling rate (see Eq. (6)).

In Eq. (44), n_i represent the occupation numbers of the infinitesimal cells $d\mathbf{r}_i$ in which the VVDS volume can be partitioned ($n_i = 0$ or 1) and the sum is intended over all the cells of the survey volume. The Fourier transform of the smoothing window function F with which the discontinuous galaxy density field is regularized is, in the case of a Top-Hat spherical smoothing filter,

$$\tilde{F}_k = 3 \left[\frac{\sin(kR)}{(kR)^3} - \frac{\cos(kR)}{(kR)^2} \right]. \quad (45)$$

In order to obtain an estimate of the quantity $\langle \tilde{\delta}_T^2(\mathbf{k}) \rangle$ that enters the Wiener filter definition (Eq. (42)), we compute the ensemble average of the squares of the modulus of the Fourier

amplitudes (Eq. (44)) and obtain

$$\begin{aligned} \langle \tilde{\delta}_T(\mathbf{k}) \tilde{\delta}_T^*(\mathbf{k}') \rangle &= \frac{1}{(\bar{\rho}V)^2} \left[\sum_i \sum_j \langle n_i n_j \rangle e^{i(\mathbf{k} \cdot \mathbf{r}_i - \mathbf{k}' \cdot \mathbf{r}_j)} \tilde{F}_k \tilde{F}_{k'}^* \right. \\ &\quad + W_k W_{k'}^* - \bar{\rho}V \sum_i \langle n_i \rangle e^{i\mathbf{k} \cdot \mathbf{r}_i} \tilde{F}_k W_{k'}^* \\ &\quad \left. - \bar{\rho}V \sum_j \langle n_j \rangle e^{i\mathbf{k}' \cdot \mathbf{r}_j} \tilde{F}_{k'}^* W_k \right] \end{aligned}$$

where

$$W(\mathbf{k}) = \frac{1}{V} \int e^{i\mathbf{k} \cdot \mathbf{r}} d^3\mathbf{r}. \quad (46)$$

The mean value of the occupation number is given by (Peebles 1980)

$$\langle n_i n_j \rangle = \bar{\rho}^2 d^3\mathbf{r}_i d^3\mathbf{r}_j [1 + \xi(\mathbf{r}_i - \mathbf{r}_j)]$$

$$\langle n_i \rangle^2 = \langle n_i \rangle = \bar{\rho} d^3\mathbf{r}_i$$

where the correlation function may be expressed via the Fourier conjugates

$$\xi(\mathbf{r}_{ij}) = \frac{1}{(2\pi)^3} \int d^3\mathbf{k} P(\mathbf{k}) e^{-i\mathbf{k} \cdot \mathbf{r}_{ij}} \quad (47)$$

$$P(\mathbf{k}) = \int d^3\mathbf{r} \xi(\mathbf{r}) e^{i\mathbf{k} \cdot \mathbf{r}}. \quad (48)$$

We convert the sums into integrals over the occupation cells taking into account the specific VVDS geometry and obtain

$$\langle \tilde{\delta}_T^2(\mathbf{k}) \rangle = |\tilde{F}_k|^2 \int d^3\mathbf{k}' P(\mathbf{k}') G(\mathbf{k} - \mathbf{k}') \quad (49)$$

$$+ |\tilde{F}_k|^2 (\bar{\rho}V)^{-1} + |W_k|^2 (1 - \tilde{F}_k)^2 \quad (50)$$

where

$$G(\mathbf{k} - \mathbf{k}') = \frac{1}{(2\pi)^3} |W(\mathbf{k} - \mathbf{k}')|^2. \quad (51)$$

Note that, in the idealized conditions of a survey of infinite spatial extension ($V \rightarrow \infty$) and no smoothing applied to data the previous expression reduces to

$$\langle \tilde{\delta}_T^2(\mathbf{k}) \rangle = (2\pi)^3 P(\mathbf{k}') \delta^D(\mathbf{k} - \mathbf{k}'). \quad (52)$$

Given the quasi pencil-beam nature of the first season VVDS data, however, the sample of the δ field can be described, to a good approximation, as confined to a cylinder of length L (aligned along the redshift direction) and radius R . In this case the Fourier transform of the survey window function (cf. Eq. (46)) is given by

$$W_k = j_0\left(\frac{L}{2}k_{\parallel}\right) 2 \frac{J_1(k_{\perp}R)}{k_{\perp}R} \quad (53)$$

where j_0 and J_1 are the spherical and first kind Bessel functions and where $k = \sqrt{k_{\parallel}^2 + k_{\perp}^2}$.

The convolution integral on the right hand side of Eq. (50) can then be evaluated as follows (e.g., Kaiser & Peacock 1991; Fisher et al. 1993)

$$\int d^3\mathbf{k}' P(\mathbf{k}') G(\mathbf{k} - \mathbf{k}') = \frac{2}{\pi V} \int_{-\infty}^{\infty} dy j_0^2\left(y - k \frac{L}{2}\right) F(y)$$

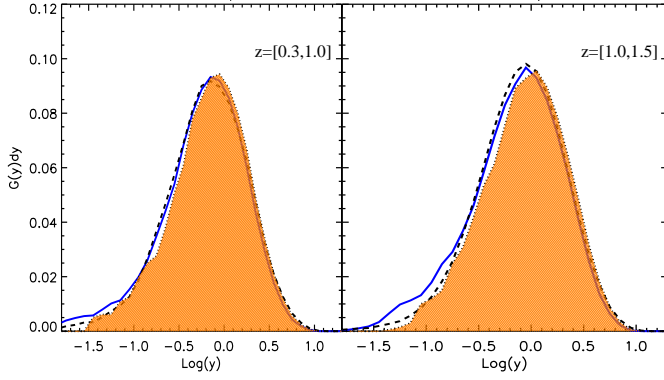


Fig. A.1. The overdensity distribution reconstructed on a scale $R = 8 h^{-1}\text{Mpc}$ and in two different redshift intervals using the GtALICS semi-analytical simulation (see Sect. 4.1) is plotted as a function of $\log(1 + \delta_g)$. The dashed line represent the distribution of overdensities traced by simulated galaxies with $M < -15 + 5 \log h$. The shaded area (dotted line) represents the *observed* overdensity distribution i.e., the distribution recovered after applying to the simulation the VVDS flux limit ($I \leq 24$) and all the underlying VVDS instrumental selection effects (see Sect. 4.1). The solid line represent the distribution of density contrasts after correcting the *observed* distribution with the Wiener technique.

where $V = \pi R^2 L$ is the volume of the cylinder and where

$$F(x) = \int_0^\infty P\left(\frac{\sqrt{x^2 + x'^2}}{R}\right) \frac{J_1^2(x')}{x'} dx'. \quad (54)$$

Note that the theoretically expected variance of the density field in the VVDS volume is derived in real comoving space. Thus, we have corrected the theoretical predictions (cf. Eq. (50)) in order to take into account redshift space distortions induced by galaxy peculiar velocities (see discussion in Sect. 5).

We compute the galaxy density field on a regular Cartesian grid of spacing $0.5 h^{-1}\text{Mpc}$ using the smoothing scheme presented in Eq. (5). The resulting 3D density map is then Fourier transformed in redshift slices having line-of-sight dimensions $dz = 0.1$. This partition strategy is implemented in order to describe consistently, using cylindrical approximations, the deep survey volume of the VVDS (whose comoving transversal dimensions are an increasing function of distance.) The Wiener filter at each wave-vector position is then computed by using Eq. (50). Finally, as described in Sect. 3, we select only the Wiener filtered density fluctuations recovered in spheres having at least 70% of their volume in the 4-passes, VVDS-02h-4 field.

In Fig. A.1 we use the GALICS semi-analytical simulation (see Sect. 4.1), to show the effect of the Wiener filter on the reconstructed galaxy density field. As Fig. A.1 shows, the net effect of the correction is to shift towards low-values the density contrasts having low signal-to-noise ratio (by definition F is always smaller than unity). Figure A.1 shows that i) at the same density, the effects of the correction are bigger at high redshift where the density field is noisier due to the increasing sparseness of a flux-limited sample, and ii) at the same redshift the Wiener filter mostly affects the low-density tail of the distribution where the counts within the TH window are small. It is also evident from Fig. A.1 that, in the density interval $-1 < \log(1 + \delta_g) < 1$, the Wiener filtered distribution offers a better approximation of the underlying PDF, than the observed (uncorrected) overdensity distribution.

Chapitre 6

Évolution de l'agrégation des galaxies depuis $z \sim 2$: une nouvelle référence

Résumé

6.1 Référence observationnelle	199
6.2 Implication sur l'évolution des galaxies	208

Ce chapitre résume les principaux résultats obtenus au cours de cette thèse. Ces travaux fournissent à la communauté scientifique internationale une nouvelle référence observationnelle quant aux propriétés d'agrégation des galaxies et leur évolution. La deuxième partie de ce chapitre présente les nouvelles contraintes apportées sur les modèles de formation et d'évolution des galaxies.

6.1 Référence observationnelle

Le *VIMOS-VLT Deep Survey* est un sondage profond de l'Univers. La première époque d'observation a fourni l'information multi-couleur et spectroscopique de ~ 11000 galaxies jusqu'à $z \sim 4 - 5$. Sa fonction de sélection, $I_{AB} = 24$ dans sa partie *Deep*, est la plus simple possible et n'introduit donc pas de biais supplémentaire sur la sélection de telle ou telle population de galaxies, contrairement à des sondages sélectionnant les objets selon des critères de couleur. De par le volume d'Univers sondé, le VVDS est le pendant à grand *redshift* des grands sondages spectro-photométriques de l'Univers local (2dFGRS, SDSS) et qui ont établi des références sur la distribution des galaxies notamment.

Les données du VVDS ont permis la mesure de l'évolution de la longueur de corrélation r_0 , pour différentes populations de galaxies, sur un très grand domaine en *redshifts* de $z \sim 0$ à $z \sim 2$, et ceci à partir du même sondage de galaxies. La longueur de corrélation de la population globale augmente ainsi lentement avec le *redshift* de $r_0 \sim 2.2 \text{ h}^{-1} \text{ Mpc}$ à $z \sim 0.3$ jusqu'à $r_0 \sim 3.6 \text{ h}^{-1} \text{ Mpc}$ pour l'intervalle en *redshift* $z = [1.3 - 2.1]$ (figure 6.1). En appliquant la fonction de sélection du sondage DEEP2, les mesures réalisées entre $z=0.7$ et $z=1.35$ sont en excellent accord avec celles de [Coil et al. \(2004\)](#). On notera cependant que DEEP2 n'est pas représentatif de la population globale des galaxies à ces

redshifts de par la présélection en couleur d'une sous population plus rouge de galaxies. Toutes les mesures réalisées dans les champs profonds VVDS-02h et VVDS-CDFS à partir des données de la première époque du VVDS sont rassemblées dans la table 6.1, les incertitudes de mesure étant quantifiées à partir de 50 champs indépendants produits à partir des simulations *GallCS*.

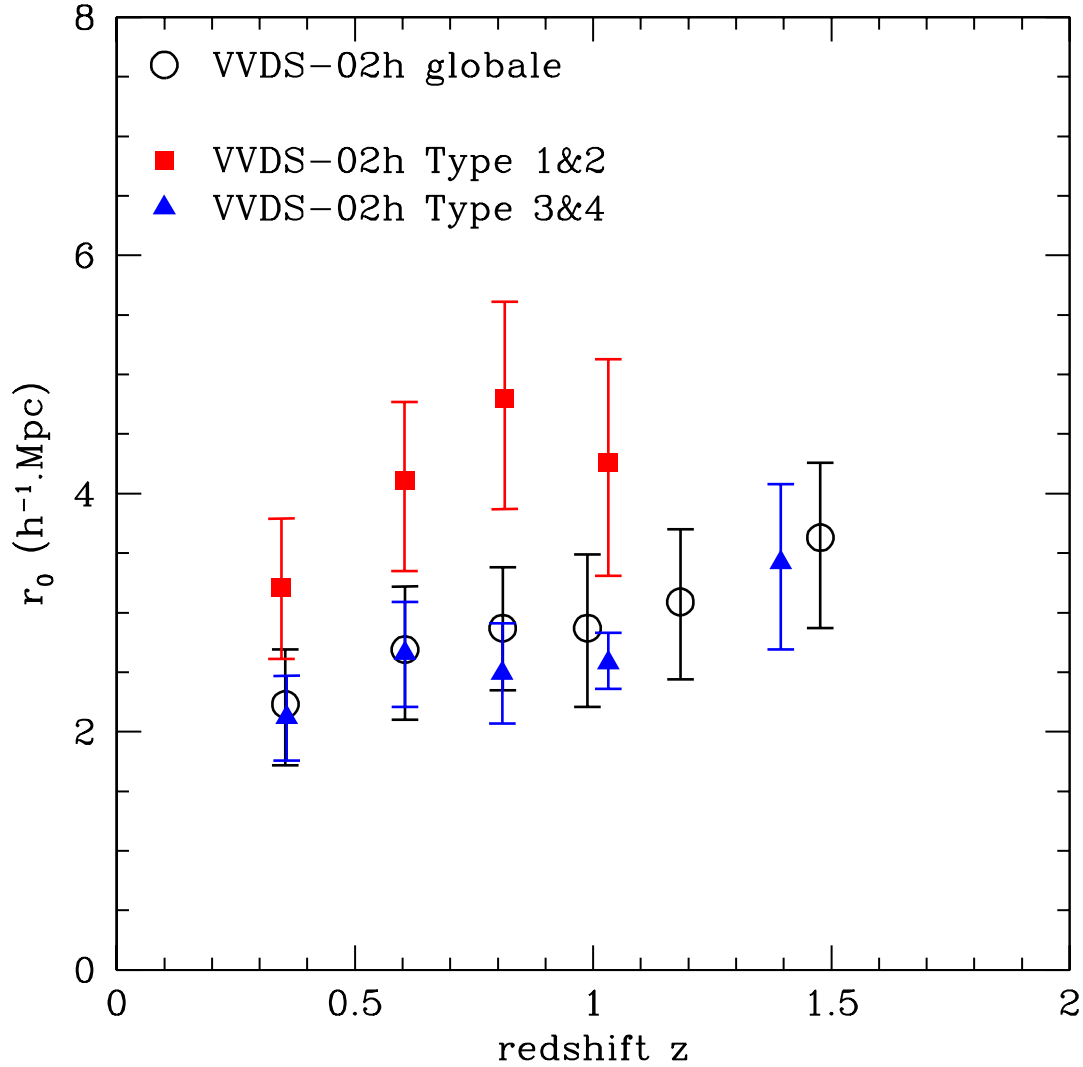


FIG. 6.1 – Évolution de la longueur de corrélation $r_0(z)$ de la population globale, de la population de type 1+2 et de la population de type 3+4 dans le VVDS-02h.

La qualité des données VVDS a ensuite permis de diviser l'échantillon global des galaxies du champ VVDS-02h selon leur type spectral et ensuite leur couleur. La classification spectrale a été réalisée sur la base des couleurs observées de chaque objets et de leur *redshift*. Les ~ 6500 galaxies dont la mesure du *redshift* est sûre à plus de 75% (note de qualité comprise entre 2 et 9) ont été réparties en 4 classes, des galaxies les plus précoces (T1) au plus tardives (T4). Les paramètres de la fonction de corrélation spatiale à

deux points ont été calculés pour différents intervalles en *redshifts* jusqu'à $z \sim 1.2$ pour les galaxies précoces et $z \sim 2$ pour les galaxies tardives. La figure 6.1 montre l'évolution des populations de types T1&2 et T3&4, comparée à celle de la population globale. La figure 6.2 indique l'évolution de chaque type spectral. On observe que quelque soit le *redshift*, les galaxies précoces sont plus fortement agrégées que les galaxies tardives. Cette ségrégation observée jusqu'à présent dans l'Univers local est maintenant contrainte pour la première fois de $z \sim 0$ jusqu'à $z \sim 1 - 2$. Toutes les mesures obtenues pour chaque type spectral de galaxies sont répertoriées dans la table 6.2.

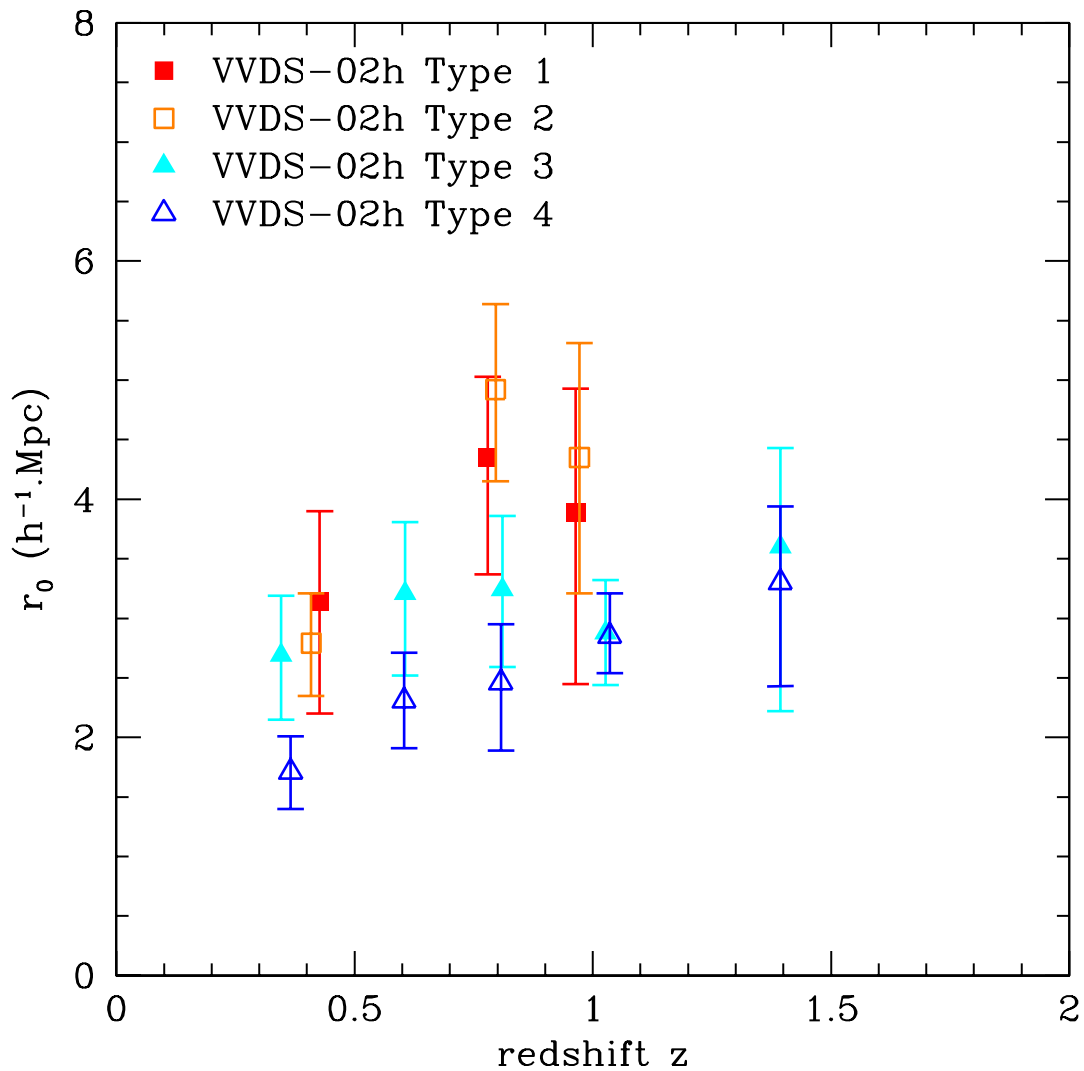


FIG. 6.2 – Évolution de la longueur de corrélation $r_0(z)$ des quatre types de populations définies dans le VVDS-02h

Les données du champs VVDS-02h présentant une distribution bimodale en couleur, l'échantillon de galaxies a alors été divisé en une population rouge et une population bleue. La figure 6.3 montre l'évolution de la longueur de corrélation r_0 mesurée pour ces deux

populations. Cette bimodalité, observée dans l'Univers local et à grand *redshift*, est ainsi étudiée sur une très grande plage de temps cosmique. La ségrégation entre les deux populations est ainsi quantifiée depuis $z \sim 1$. Les différentes valeurs mesurées sont résumées dans le tableau 6.3.

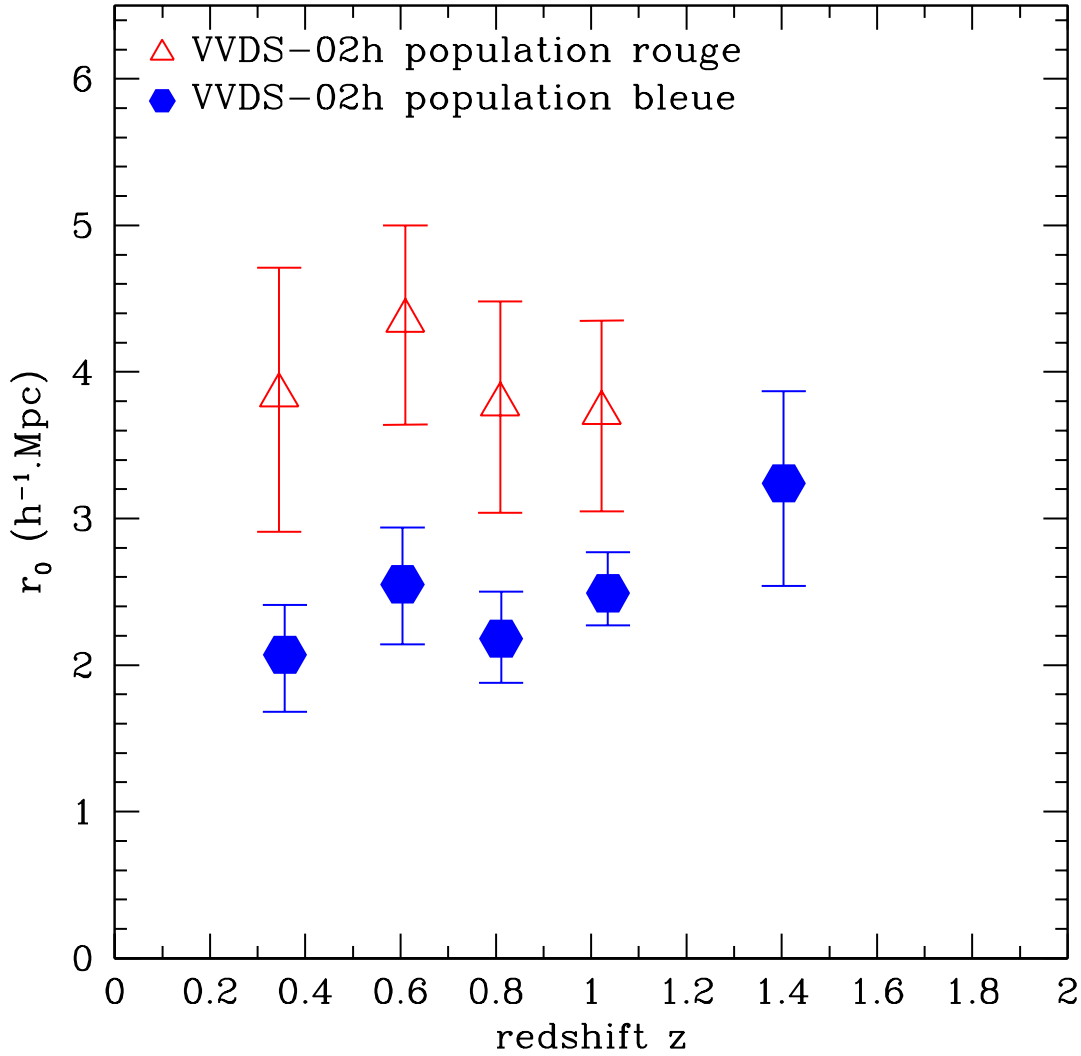


FIG. 6.3 – Évolution de la longueur de corrélation $r_0(z)$ des populations rouges et bleues définies dans le VVDS-02h

Enfin, la combinaison des observations photométriques réalisées dans l'ultraviolet par le satellite GALEX et les données spectroscopiques du VVDS ont permis une première détermination des paramètres r_0 et γ de populations de galaxies présentant une formation stellaire importante, jusqu'à $z \sim 1$ en NUV et $z \sim 0.5$ en FUV. La figure 6.4 décrit l'évolution observée en NUV et FUV. Les valeurs sont rassemblées dans la table 6.4.

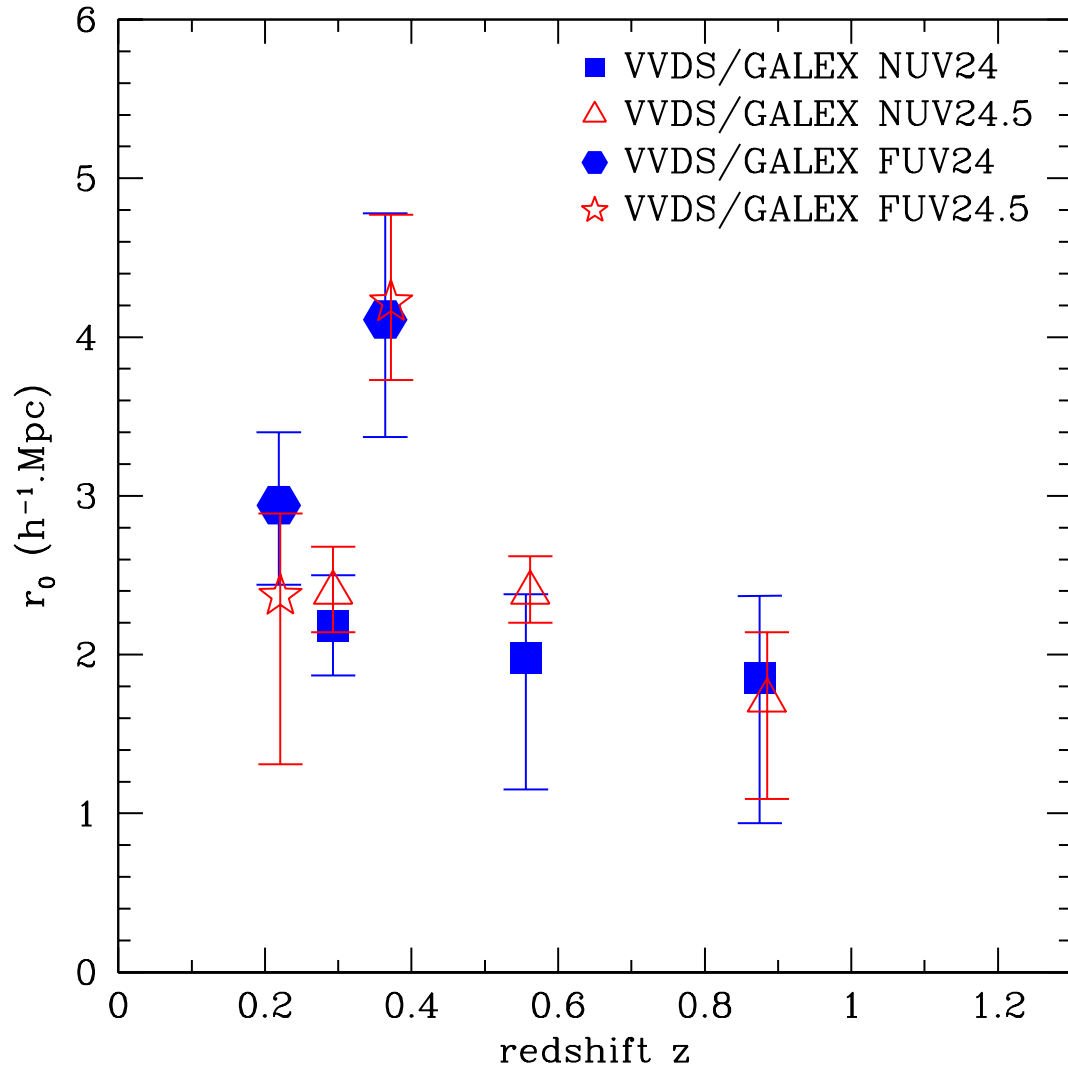


FIG. 6.4 – Évolution de la longueur de corrélation $r_0(z)$ des populations sélectionnées en NUV et FUV dans le VVDS-02h

Echantillon	Nombre de galaxies	Intervalle en <i>redshift</i>	<i>redshift</i> moyen	$M_B - 5\log(h)$ (moyenne)	$(B - I)_{(z=0)}$	r_0 ($h^{-1} \cdot \text{Mpc}$)	γ	r_0 avec γ fixé ($h^{-1} \cdot \text{Mpc}$)
VVDS-02h	1258	0.2 – 0.5	0.354	−17.57	1.34	$2.23^{+0.46}_{-0.51}$	$1.67^{+0.17}_{-0.11}$	$2.27^{+0.47}_{-0.52}$ ($\gamma=1.76$)
VVDS-02h	1443	0.5 – 0.7	0.605	−18.68	1.28	$2.69^{+0.53}_{-0.59}$	$1.71^{+0.18}_{-0.11}$	$2.51^{+0.49}_{-0.55}$
VVDS-02h	1398	0.7 – 0.9	0.810	−19.20	1.36	$2.87^{+0.51}_{-0.52}$	$1.72^{+0.19}_{-0.12}$	$2.87^{+0.50}_{-0.52}$
VVDS-02h	1025	0.9 – 1.1	0.988	−19.66	1.35	$2.87^{+0.62}_{-0.66}$	$1.77^{+0.24}_{-0.14}$	$2.75^{+0.60}_{-0.63}$
VVDS-02h	561	1.1 – 1.3	1.183	−20.26	1.35	$3.09^{+0.61}_{-0.65}$	$1.96^{+0.27}_{-0.21}$	$2.93^{+0.58}_{-0.62}$
VVDS-02h	432	1.3 – 2.1	1.477	−20.93	1.25	$3.63^{+0.63}_{-0.76}$	$1.92^{+0.31}_{-0.31}$	$3.69^{+0.77}_{-1.00}$
VVDS-02h “DEEP2”	1620	0.7 – 1.35	0.966	−20.00	1.36	$3.05^{+0.51}_{-0.53}$	$1.56^{+0.14}_{-0.11}$	$3.17^{+0.53}_{-0.55}$ ($\gamma=1.66$)
VVDS-02h “DEEP2”	687	0.7 – 0.9	0.827	−19.65	1.41	$3.59^{+0.77}_{-0.83}$	$1.65^{+0.25}_{-0.17}$	$3.59^{+0.77}_{-0.83}$
VVDS-02h “DEEP2”	933	0.9 – 1.35	1.068	−20.25	1.32	$3.29^{+0.57}_{-0.59}$	$1.82^{+0.19}_{-0.15}$	$3.13^{+0.54}_{-0.57}$
VVDS-CDFS	421	0.4 – 0.7	0.584	−19.44	1.30	$3.19^{+0.93}_{-1.01}$	$1.45^{+0.25}_{-0.17}$	$3.31^{+1.06}_{-1.26}$ ($\gamma=1.76$)
VVDS-CDFS	372	0.6 – 0.8	0.693	−19.74	1.47	$4.55^{+1.25}_{-1.46}$	$1.48^{+0.28}_{-0.15}$	$4.15^{+1.25}_{-1.50}$
VVDS-CDFS	452	0.8 – 1.5	1.060	−20.43	1.34	$3.53^{+1.29}_{-1.66}$	$1.66^{+0.40}_{-0.27}$	$3.49^{+1.28}_{-1.78}$

TAB. 6.1 – Tableau récapitulatif des valeurs des longueurs de corrélation r_0 et pentes γ mesurés dans le VVDS-02h et VVDS-CDFS pour la population globale au cours de cette thèse

Echantillon	Nombre de galaxies	Intervalle en <i>redshift</i>	<i>redshift</i> moyen	$M_{\text{BAB}} - 5\log(h)$ moyennne	$5\log(h)$ médianne	r_0 ($h^{-1} \cdot \text{Mpc}$)	γ	r_0 avec γ fixé ($h^{-1} \cdot \text{Mpc}$)
Type 1	164	0.2 – 0.6	0.427	−18.80	−19.02	$3.14^{+0.76}_{-0.94}$	$2.41^{+0.33}_{-0.37}$	$3.51^{+0.65}_{-0.71}$ ($\gamma=2.15$)
	332	0.6 – 1.0	0.779	−19.76	−19.82	$4.35^{+0.68}_{-0.98}$	$2.01^{+0.46}_{-0.29}$	$3.93^{+0.55}_{-0.74}$
	237	0.8 – 1.2	0.964	−20.32	−20.35	$3.89^{+1.04}_{-1.44}$	$2.04^{+0.80}_{-0.45}$	$3.62^{+0.89}_{-1.26}$
Type 2	299	0.2 – 0.6	0.409	−18.35	−18.42	$2.79^{+0.42}_{-0.44}$	$1.88^{+0.17}_{-0.14}$	$2.76^{+0.35}_{-0.32}$ ($\gamma=1.90$)
	420	0.6 – 1.0	0.796	−19.53	−19.50	$4.92^{+0.72}_{-0.77}$	$1.82^{+0.27}_{-0.20}$	$4.89^{+0.58}_{-0.58}$
	358	0.8 – 1.2	0.972	−20.06	−20.12	$4.35^{+0.96}_{-1.14}$	$2.00^{+0.38}_{-0.26}$	$4.62^{+0.85}_{-0.96}$
Type 3	437	0.2 – 0.5	0.346	−17.48	−17.33	$2.69^{+0.50}_{-0.54}$	$1.83^{+0.16}_{-0.11}$	$2.73^{+0.44}_{-0.42}$ ($\gamma=1.78$)
	470	0.5 – 0.7	0.606	−18.67	−18.60	$3.21^{+0.60}_{-0.69}$	$1.91^{+0.24}_{-0.15}$	$3.33^{+0.54}_{-0.59}$
	473	0.7 – 0.9	0.810	−19.30	−19.21	$3.24^{+0.62}_{-0.65}$	$1.56^{+0.22}_{-0.15}$	$3.21^{+0.60}_{-0.60}$
	439	0.9 – 1.2	1.027	−19.84	−19.84	$2.88^{+0.44}_{-0.44}$	$1.83^{+0.36}_{-0.27}$	$2.82^{+0.49}_{-0.50}$
	191	1.2 – 2.0	1.394	−20.76	−20.75	$3.60^{+0.83}_{-1.38}$	$1.73^{+0.35}_{-0.46}$	$3.65^{+1.02}_{-3.04}$
Type 4	507	0.2 – 0.5	0.366	−17.26	−17.17	$1.71^{+0.30}_{-0.31}$	$1.80^{+0.16}_{-0.12}$	$1.71^{+0.30}_{-0.30}$ ($\gamma=1.86$)
	598	0.5 – 0.7	0.604	−18.28	−18.14	$2.31^{+0.40}_{-0.40}$	$1.76^{+0.14}_{-0.10}$	$2.28^{+0.42}_{-0.39}$
	534	0.7 – 0.9	0.807	−18.94	−18.83	$2.46^{+0.49}_{-0.57}$	$1.92^{+0.31}_{-0.19}$	$2.49^{+0.53}_{-0.58}$
	591	0.9 – 1.2	1.036	−19.55	−19.49	$2.85^{+0.36}_{-0.31}$	$1.94^{+0.18}_{-0.15}$	$2.91^{+0.40}_{-0.37}$
	289	1.2 – 2.0	1.394	−20.44	−20.40	$3.30^{+0.64}_{-0.87}$	$1.92^{+0.36}_{-0.32}$	$3.27^{+1.01}_{-2.35}$
Type 1&2	307	0.2 – 0.5	0.346	−18.33	−18.51	$3.21^{+0.58}_{-0.60}$	$1.94^{+0.27}_{-0.29}$	$3.21^{+0.40}_{-0.40}$ ($\gamma=1.89$)
	373	0.5 – 0.7	0.604	−19.00	−19.02	$4.11^{+0.66}_{-0.76}$	$1.94^{+0.19}_{-0.12}$	$4.14^{+0.53}_{-0.57}$
	390	0.7 – 0.9	0.814	−19.72	−19.64	$4.80^{+0.81}_{-0.93}$	$1.76^{+0.30}_{-0.18}$	$4.41^{+0.62}_{-0.69}$
	371	0.9 – 1.2	1.032	−20.33	−20.38	$4.26^{+0.87}_{-0.95}$	$1.92^{+0.31}_{-0.24}$	$4.32^{+0.79}_{-0.90}$
Type 3&4	943	0.2 – 0.5	0.357	−17.36	−17.23	$2.12^{+0.35}_{-0.36}$	$1.78^{+0.14}_{-0.09}$	$2.12^{+0.32}_{-0.32}$ ($\gamma=1.76$)
	1068	0.5 – 0.7	0.605	−18.45	−18.32	$2.66^{+0.43}_{-0.45}$	$1.74^{+0.15}_{-0.09}$	$2.66^{+0.41}_{-0.41}$
	1007	0.7 – 0.9	0.809	−19.11	−19.01	$2.49^{+0.42}_{-0.42}$	$1.67^{+0.17}_{-0.11}$	$2.46^{+0.48}_{-0.50}$
	1030	0.9 – 1.2	1.032	−19.68	−19.63	$2.58^{+0.25}_{-0.22}$	$1.86^{+0.11}_{-0.08}$	$2.58^{+0.28}_{-0.25}$
	480	1.2 – 2.0	1.394	−20.57	−20.51	$3.42^{+0.66}_{-0.73}$	$1.99^{+0.24}_{-0.20}$	$3.24^{+0.90}_{-1.23}$

TAB. 6.2 – Tableau récapitulatif des valeurs des longueurs de corrélation r_0 et pentes γ mesurés dans le VVDS-02h en divisant la population globale par type spectral au cours de cette thèse

Population de galaxies	Intervalle en <i>redshift</i>	<i>redshift</i> moyen	Nombre de galaxies ($\langle M_{B_{AB}} - 5\log(h) \rangle$)			$r_0(z)$ ($h^{-1} \cdot \text{Mpc}$)	γ
			Total	dont T1&2	dont T3&4		
rouge	0.2 – 0.5	0.345	270 (−17.86)	221 (−18.18)	49 (−16.40)	$3.84^{+0.87}_{-0.93}$	$1.88^{+0.23}_{-0.16}$
	0.5 – 0.7	0.610	277 (−18.79)	253 (−18.89)	24 (−17.65)	$4.35^{+0.65}_{-0.71}$	$1.99^{+0.22}_{-0.18}$
	0.7 – 0.9	0.809	355 (−19.57)	313 (−19.71)	42 (−18.47)	$3.78^{+0.70}_{-0.74}$	$1.87^{+0.28}_{-0.22}$
	0.9 – 1.2	1.022	296 (−20.22)	283 (−20.26)	13 (−19.31)	$3.72^{+0.63}_{-0.67}$	$1.83^{+0.32}_{-0.32}$
bleue	0.2 – 0.5	0.357	981 (−17.52)	86 (−18.70)	895 (−17.41)	$2.07^{+0.34}_{-0.39}$	$1.67^{+0.14}_{-0.09}$
	0.5 – 0.7	0.603	1165 (−18.54)	121 (−19.22)	1044 (−18.47)	$2.55^{+0.39}_{-0.41}$	$1.68^{+0.15}_{-0.09}$
	0.7 – 0.9	0.811	1042 (−19.18)	77 (−19.75)	965 (−19.14)	$2.18^{+0.32}_{-0.30}$	$1.40^{+0.16}_{-0.11}$
	0.9 – 1.2	1.035	1104 (−19.75)	88 (−20.56)	1016 (−19.68)	$2.49^{+0.28}_{-0.22}$	$1.84^{+0.14}_{-0.10}$
	1.2 – 2.0	1.404	574 (−20.65)	94 (−21.08)	480 (−20.57)	$3.24^{+0.63}_{-0.70}$	$1.91^{+0.20}_{-0.18}$

TAB. 6.3 – Tableau récapitulatif des valeurs des longueurs de corrélation r_0 et pentes γ mesurés dans le VVDS-02h pour les populations rouges et bleues au cours de cette thèse

Echantillon	Nombre de galaxies			Intervalle en <i>redshift</i>	<i>redshift</i> moyen	$M - 5\log(h)$			r_0 ($h^{-1} \cdot \text{Mpc}$)	γ
	total	T1&2	T3&4			B_{AB}	NUV_{AB}	FUV_{AB}		
$NUV \leq 24$	223	47	176	0.2 – 0.4	0.293	-18.27	-16.84	-16.60	$2.18^{+0.32}_{-0.31}$	$2.08^{+0.29}_{-0.24}$
	263	18	245	0.4 – 0.7	0.556	-19.41	-17.95	-17.98	$1.98^{+0.40}_{-0.83}$	$1.55^{+0.28}_{-0.31}$
	154	2	152	0.7 – 1.2	0.875	-20.15	-18.93	-19.01	$1.85^{+0.52}_{-0.91}$	$2.43^{+1.73}_{-0.63}$
$NUV \leq 24.5$	320	67	253	0.2 – 0.4	0.293	-17.98	-16.52	-16.28	$2.39^{+0.29}_{-0.25}$	$2.10^{+0.22}_{-0.19}$
	428	36	392	0.4 – 0.7	0.562	-19.19	-17.69	-17.62	$2.39^{+0.23}_{-0.19}$	$2.09^{+0.16}_{-0.15}$
	269	4	265	0.7 – 1.2	0.885	-19.97	-18.77	-18.69	$1.71^{+0.43}_{-0.62}$	$2.64^{+1.00}_{-0.45}$
$FUV \leq 24$	128	26	102	0.1 – 0.3	0.219	-17.84	-16.40	-16.20	$2.94^{+0.46}_{-0.50}$	$1.55^{+0.20}_{-0.17}$
	124	13	111	0.3 – 0.5	0.364	-18.84	-17.41	-17.34	$4.11^{+0.67}_{-0.74}$	$1.85^{+0.32}_{-0.24}$
$FUV \leq 24.5$	179	39	140	0.1 – 0.3	0.221	-17.59	-16.12	-15.90	$2.37^{+0.52}_{-1.06}$	$1.36^{+0.24}_{-0.20}$
	205	22	183	0.3 – 0.5	0.372	-18.56	-17.13	-17.04	$4.22^{+0.55}_{-0.49}$	$1.80^{+0.14}_{-0.12}$

TAB. 6.4 – Tableau récapitulatif des valeurs des longueurs de corrélation r_0 et pentes γ mesurés dans le VVDS-02h à partir des échantillons sélectionnés en ultraviolet avec le satellite GALEX au cours de cette thèse

6.2 Implication sur l'évolution des galaxies

Les mesures de l'agrégation des galaxies réalisées dans le sondage VVDS établissent maintenant de nouvelles références en la matière. Ces résultats apportent une vision détaillée de l'évolution de la distribution des galaxies au cours du temps qui doit maintenant être prise en compte par les modèles d'évolution.

L'interprétation physique des résultats est complexe. La fonction de sélection du VVDS, purement en magnitude, implique que la luminosité absolue des populations sondées augmente avec le *redshift*. Or il est important de comparer des mesures réalisées pour des populations similaires. Ainsi, l'amplitude de la fonction de corrélation des galaxies du VVDS détectées à $z \sim 1 - 2$ était ~ 2.5 fois plus faible que celle observée localement dans le 2dFGRS pour des galaxies de même luminosité.

La lente variation de la longueur de corrélation de la population globale à mesure que le *redshift* augmente est différente de celle prédite par les modèles d'évolution de la matière noire. Nos observations mettent en évidence une forte évolution du biais qui existe entre la distribution spatiale des galaxies et celle de la matière noire, en accord avec les simulations prenant en compte l'évolution des structures et les processus physiques internes aux galaxies. Les galaxies très lumineuses à $z \sim 1 - 2$ traceraient alors uniquement les pics principaux du champs de densité de la matière noire.

La ségrégation observée entre les galaxies de type précoce et celles de type tardif est déjà très fortement marquée à $z \sim 1$. De même, les galaxies rouges sont plus fortement agrégées que les galaxies bleues. La forte amplitude de la fonction de corrélation des galaxies de type précoce à $z \sim 1$ semble donc indiquer que celles-ci sont déjà en place au sein des structures voici ~ 7 à 9 milliards d'années. L'agrégation des galaxies tardives continuent par contre à évoluer et la formation stellaire se déplace des régions denses vers les régions moins denses. Ces résultats sont largement en accord avec un modèle hiérarchique de formation des galaxies. Le biais relatif entre les galaxies précoces et tardives évolue en effet très peu entre $z \sim 1$ et $z=0$, ainsi que prédisent les simulations numériques (Kauffmann *et al.*, 1999a).

Conclusions et perspectives

Conclusion ...

Ma thèse s'est déroulée au sein du consortium franco-italien du *VIMOS-VLT Deep Survey* (VVDS). Le VVDS est un relevé spectroscopique profond associé à un relevé photométrique multi-couleur. Un des objectifs majeurs du VVDS consiste à retracer l'évolution des grandes structures observées dans l'Univers jusqu'à une époque où celui-ci n'était âgé que de $\sim 10\%$ de son âge actuel.

Au sein de cette équipe internationale, j'ai participé à la réduction des données spectroscopiques et à la mesure de plus de 6000 *redshifts*. En parallèle, j'ai particulièrement travaillé sur la mesure de la fonction de corrélation spatiale, et sur l'estimation des erreurs de mesure. J'ai réalisé des catalogues fictifs de galaxies reproduisant la géométrie des champs profonds du VVDS ainsi que la stratégie d'observation employée avec ses différents biais à partir des simulations *GallCS* développées à l'Institut d'Astrophysique de Paris. Ces catalogues m'ont permis de tester le code de calcul de la fonction de corrélation spatiale à deux dimensions $\xi(r_p, \pi)$ et de m'assurer de notre capacité à bien corriger les différents biais observationnels et d'incomplétude du VVDS.

J'ai écrit un ensemble de programmes permettant d'automatiser le calcul de la fonction de corrélation spatiale $\xi(r_p, \pi)$, sa projection $w_p(r_p)$, la longueur de corrélation r_0 , la pente γ et leurs erreurs associées. Cette distribution et les différents catalogues fictifs ont été mis à la disposition du consortium et accompagnés d'une documentation d'utilisation.

Cet ensemble de programme m'a permis de mesurer la fonction de corrélation spatiale et son évolution depuis *redshift* $z \sim 2$ à partir des données spectroscopiques des champs profonds VVDS-02h et VVDS-CDFS, pour la population globale de galaxies, puis divisée en types spectraux, en couleur, et pour un sous échantillon sélectionné en ultraviolet. C'est la première fois que de telles mesures sont réalisées sur une si grande plage en *redshift*. Elles font maintenant office de nouvelles références dans l'étude des grandes structures présentes dans l'Univers.

... et perspectives

Les outils que j'ai développés vont être utilisés pour conduire plusieurs nouvelles analyses des données spectroscopiques du sondage VVDS.

Tout d'abord, dans la partie *Deep*, la mesure de l'agrégation des galaxies va être réalisée pour un échantillon limité en volume, où les galaxies sont sélectionnées en fonction de leur luminosité intrinsèque. Il sera possible de quantifier la ségrégation en luminosité observée dans l'Univers local sur une très grande plage de temps cosmique et de relier cette étude à celles présentées dans cette thèse. Les données spectroscopiques du VVDS

permettent aussi une mesure de l'agrégation des galaxies à très grand décalage spectral $z \sim 3 - 5$ et de comparer les résultats alors obtenus aux observations déjà réalisées à ces *redshifts*, mais à partir de fonctions de sélection plus restrictives. On observe en effet dans le VVDS une densité de galaxies supérieure d'un facteur 2 à 6 à ces *redshifts* que ce qui était jusqu'à présent établi, remettant ainsi en cause les modèles de formation et d'évolution des galaxies (Le Fèvre *et al.*, 2005b).

Enfin, le champ VVDS-02h a été observé en infrarouge avec le satellite Spitzer, et dans le cadre d'une collaboration avec le consortium SWIRE, une mesure de la fonction de corrélation des objets sélectionné dans ces grandes longueurs d'onde est prévue. L'exploitation des données infrarouge permettra d'étudier la distribution des galaxies en fonction de leur masse et son évolution.

Jusqu'à présent, les données de la seconde époque, et en particulier celles des champs *Wide* n'ont pas encore été exploitées. Leur utilisation permettra de mieux contraindre la variance cosmique et d'étudier les structures sur de très grandes échelles, de l'ordre de $100 h^{-1} \text{Mpc}$ à $z \sim 1$.

Il est aussi prévu une exploitation des *redshifts* photométriques, maintenant relativement bien contraints grâce à la très large couverture multi-couleur du champs VVDS-02h, de l'ultraviolet avec GALEX, jusqu'à l'infrarouge lointain avec Spitzer. Ces données élargiront donc la taille des échantillons de galaxies à des millions d'objets.

A moyen terme, le relevé COSMOS permettra d'allier l'information spectroscopique de VIMOS avec l'information morphologique obtenue à partir des images de la caméra à haute résolution ACS du télescope spatial Hubble. Il sera alors possible d'identifier jusqu'à $z \sim 1$ le type morphologique de chaque galaxie d'un échantillon de taille sans précédent. La mesure de la fonction de corrélation de galaxies détectées en bande H permettra d'étudier la distribution de la masse stellaire, et des analyses similaires à celles présentées ici permettront là encore de bien estimer la variance cosmique. Dans l'avenir, les observations avec le futur télescope spatial *James Webb Space Telescope* (JWST) nous révéleront la distribution et les propriétés physiques des premières galaxies, situées à très grand *redshift*, à une période où l'Univers était âgé de quelques millions à quelques milliards d'années. Toutes ces mesures directes de la distribution des galaxies nous permettront de retracer la séquence d'événements et de processus physiques qui ont conduit à la structuration remarquable de l'Univers.

Annexe A

Correlation Function Tool - release xi_v0.9.8

Cette annexe présente la documentation qui décrit l'utilisation des différents programmes de calcul de la fonction de corrélation spatiale $\xi(r_p, \pi)$, des paramètres r_0 et γ et leurs erreurs associées. Cette documentation accompagne la suite de programmes mise à la disposition des membres du consortium VVDS.

Introduction to CFT

The rule of this document is to help the reader to use this “Correlation Function Tool” (CFT hereafter). The idea is to describe step by step everything he/she must do in order to get, at the end, the correlation length r_0 of his/her data and the associated errors.

The part [A.1](#) explains the basic things required to use the CFT. If the reader has a problem in this part, he should contact the author¹. Indeed, the tools will not work. But, waiting for an update of the tool or more information, the user can of course read the following parts of this documentation ;-)

The part [A.2](#) describes the programs used to compute $\xi(s)$ and $\xi(r_p, \pi)$, with VVDS data and GallCS simulation.

The part [A.3](#) explains how to measure the quantities the reader is interested in (r_0 , γ , error contours, ...).

The part [A.4](#) is more technical (...) but the user will have to understand it to configure properly several files. Indeed, even if the CFT is quasi *press-button*, the user will still have to change few things by himself.

Then, the part [A.5](#) describes the different GallCS catalogues.

Have fun with CFT and exciting scientific results !!

¹baptiste.meneux@oamp.fr

A.1 Step 1 : What you need with CFT

A.1.1 Libraries

Most of the programs used within the CFT are written in Fortran. You will need libraries to read FITS file. The cfitsio routines can be found on the same web page where this file was downloaded . You can find on the official website² the last version. You must install it before using the different macros (part A.2).

A.1.2 Compiler

Fortran

The default compiler is ifc from Intel Fortran Compiler. If you do not have this compiler, you can also use any GNU f95 compiler for Linux. Run the macro called

test_compiler

found in the main directory to test if you have one of these compilers and which one will be used through the macros. If you do not have any of these compilers, send an email to the author.

If the previous test is ok, you won't have to do something special to compile the programs. It will be done automatically with the macro

compiler

found in the directories where the codes are, and called in the different macros (part A.2).

Perl

You need a Perl distribution installed on your machine. If it's not the case, you will not be able to apply a bootstrap resampling method on your data (part A.2.2). You can download a Perl distribution on the web³.

A.2 Step 2 : $\xi(s)$ and $\xi(r_p, \pi)$

The first time the CFT is applied to a set of data, in a redshift bin for a given geometry (pointings/quadrants), a random sample is built and stored in an appropriate directory. This sample can be used again, in the same condition. In this case, the computation is faster.

In this part, all the programs described are scripts written in shell. They are explained here below. Move in the directory MACRO_CODE.

It is advised to redirect the screen output of all the macros in a log file. Here is an example with a macro using 2 parameters :

macro [parameter1] [parameter2] > macro_parameter1_parameter2.log 2>&1

If the output files are not renamed within a macro, their original name is :

²see <http://heasarc.gsfc.nasa.gov/docs/software/fitsio/fitsio.html>

³see <http://www.perl.com/download.csp>

fort.11	$\xi(r_p, \pi)$	(linear , linear)
fort.12	$\xi(r_p, \pi)$	(log , log)
fort.14	$\xi(r_p, \pi)$	(log , linear)
fort.13	$\xi(s)$	(log)
fort.17	$\xi(s)$	(linear)

A.2.1 The VVDS data

To compute $\xi(s)$ and $\xi(r_p, \pi)$ from your data, you can use the macro called *macro_data_fast*. It takes 5 parameters :

macro_data_fast [Field] [z₁] [z₂] [Ω_m] [Ω_Λ]

With

- Field : F02, CDFS, DEEP2, F10, ...
- z₁ : minimum redshift of the bin
- z₂ : maximum redshift of the bin
- Ω_m
- Ω_Λ

This macro applies the “sspoc” code (see part A.4.1) to the VVDS data. It generates the random catalog with the good geometry if it does not exist yet. This random sample is then stored in a directory.

The output files are stored in

../[Field]/Result_Data_[Ω_m][Ω_Λ]/Slice_[z₁][z₂]/

A.2.2 The VVDS data and bootstrap

Now you want to compute the errors of your measured $\xi(s)$ and $\xi(r_p, \pi)$, using a bootstrap resampling method. You can use the macro called *macro_bootstrap_fast*. It takes 6 parameters :

macro_bootstrap_fast [Field] [z₁] [z₂] [Nboot] [Ω_m] [Ω_Λ]

With

- Field : F02, CDFS, DEEP2, F10, ...
- z₁ : minimum redshift of the bin
- z₂ : maximum redshift of the bin
- Nboot : number of bootstrap loop
- Ω_m
- Ω_Λ

This macro builds *Nboot* times a different input catalogue by bootstrap resampling the original input catalogue. Then the “sspoc” code (see part A.4.1) is applied to these *Nboot* input files.

Input bootstrap files are stored in

../[Field]/Input_bootstrap_files/Slice_[z₁][z₂]/

All these files can be deleted. They are only stored for check-up, if necessary. Output files are stored in `../[Field]/Result_Bootstrap_ $[\Omega_m]$ _ $[\Omega_\Lambda]$ /Slice_ $[z_1]$ _ $[z_2]$ /`

- XIRPPI_LINLIN
- XIRPPI_LOGLOG
- XIRPPI_SEMILOG
- XIS_LIN
- XIS_LOG

A.2.3 The GalICS data ⁴

In observing conditions

You have 50 catalogues from the 50 GalICS cones. These catalogues were made applying your observing strategy (magnitude selection, VIMOS selection, ...). You want to compute $\xi(s)$ and $\xi(r_p, \pi)$ in a given redshift slice for all the cones. You will use the macro called *macro_galics_fast*. It takes 3 parameters :

`macro_galics_fast [Field] $[z_1]$ $[z_2]$`

With

- Field : F02, CDFS, DEEP2, F10, ...
- z_1 : minimum redshift of the bin
- z_2 : maximum redshift of the bin

Here you can not chose the cosmology which is $\Omega_m = 0.333$ and $\Omega_\Lambda = 0.667$ for the GalICS simulation.

The macro applies the code to the GalICS data. The random catalogue is generated for the first cone and then used with all the 50 cones.

Output files are stored in `../[Field]/Result_GalICS/Slice_ $[z_1]$ _ $[z_2]$ /`

- XIRPPI_LINLIN
- XIRPPI_LOGLOG
- XIRPPI_SEMILOG
- XIS_LIN
- XIS_LOG

As a reference

You now want to compute the true underlying $\xi(s)$ and $\xi(r_p, \pi)$ without any selection effects. You can use the macro called *macro_galics_true_fast*. It takes 3 parameters :

`macro_galics_true_fast [Field] $[z_1]$ $[z_2]$`

With

- Field : F02, CDFS, DEEP2, F10, ...
- z_1 : minimum redshift of the bin
- z_2 : maximum redshift of the bin

⁴See part [A.5](#)

It applies the “true” code (see part A.4.2) that computes the true correlation function (the reference) from the GALICS files with no photometric mask, no SSPOC selection, ...

You must define the smaller area that cover all the observed field. You may not need to do this step if you do not want to know how well you are able to measure the correlation function in your observing conditions.

The random catalogue is generated for the first cone and then used with all the 50 cones.

Output files are stored in

../GALICS/Result_True/[Field]/Slice_[z1]_[z2]/

A.2.4 In practice

In the directory MACRO_CODE, there is a last shell macro called *play*. You can edit it. Then, you just have to uncomment or to add command lines to call the other macros with the good syntax.

Run *play* without any argument, in background. If everything is setup correctly, you will just have to wait for a while ! You can not begin the step 2 while this first step is not finished.

A.3 Step 3 : r_0 and γ

The different shell macro used to compute and to fit $w_p(r_p)$ are in the directory MACRO_FIT.

A.3.1 Computation and fit of $w_p(r_p)$

The name of the main macro is *macro_fit_v3*. It needs 9 parameters defined below. This macro computes the best parameters r_0 and γ for the function $w_p(r_p)$ computed from a given function $\xi(r_p, \pi)$. It uses the fortran program *fitwplog_v7.f* (see part A.4.4). This program gives in output the values of $w_p(r_p)$ and its associated best fit, the corresponding values r_0 and γ and their errors. Everything is computed with and without the covariance matrix. The error contours are written in binaries files while the values of $w_p(r_p)$ are given in an ASCII file. These output files are moved in the sub-directory FIT of your working directory [field]. The 9 parameters of this macro are :

1. z_1 , minimum redshift of the bin
2. z_2 , maximum redshift of the bin
3. number of the last r_p bin for the fit
4. π_{\max} , maximal integration value
5. – 1 \rightarrow fit $w_p(r_p)$ of VVDS data with bootstrap
 – 2 \rightarrow fit $w_p(r_p)$ of <GALICS sspoc> with GALICS sspoc
 – 3 \rightarrow fit $w_p(r_p)$ of <GALICS true> with GALICS true
 – 4 \rightarrow fit $w_p(r_p)$ of one GALICS sspoc cone with GALICS sspoc
 – 5 \rightarrow fit $w_p(r_p)$ of one GALICS true cone with GALICS true
6. field (F02, CDFS, ...)

7. number of the GallCS cone to fit if the 5th parameter is equal to 4 or 5, else 0
8. Ω_m
9. Ω_Λ

A.3.2 The slope γ

This macro, *fix_slope*, allows you to fit a given $w_p(r_p)$ with a given slope γ . The parameter of this macro is your value of γ .

If you do not give any parameter, the slope will be let free when doing the fit.

This macro must be used before *macro_fit_v3* (part A.3.1).

A.3.3 The r_p bins not used when fitting

If you do not want to use one or several r_p bin(s) to fit $w_p(r_p)$, you must use the macro

make_pts_to_be_removed.

The parameter(s) is(are) the list (with or without any order) of the r_p bin(s) you do not want to take account for.

It will have no effect if you do not give any parameters to this macro. In this case, all the r_p points (from the first to the last one, defined in parameter 3 of *macro_fit_v3*, part A.3.1) will be used by default to fit $w_p(r_p)$.

This macro must be used before *macro_fit_v3* (part A.3.1).

A.3.4 The fit of the VVDS and GallCS data

The goal of *macro_mix_fit* is to compute everything you need to measure the correlation length r_0 and the slope γ of your data, in a given redshift bin, with the error contours computed from GallCS in the same conditions. It applies *macro_fit_v3* twice with the good parameters.

Of course you must feed *macro_mix_fit* with several parameters as the limits of the redshift bin, the working directory (the field) of your data, ... You can also fix the slope and/or decide not to use such or such r_p bin to fit your data, using the previous macros (part A.3.2, A.3.3).

A.3.5 In practice

The macro *macro_show_all* is an exemple of what can be done with the previous macro. It allows you to compute in an automatic way the fit of several function $w_p(r_p)$, in different redshift bins for exemple. Feel free to update this macro by yourselves.

A.3.6 The results with SM routines

Two files can be found in the sub directory called "sm". This files contain exemple of Super Mongo routines to plot your result. You will have to adapt them to the result you want to plot. For more information, fell free to contact the author ;-)

A.4 Programs

A.4.1 To compute the observed $\xi(s)$ and $\xi(r_p, \pi)$

The main program is *xi_cosmo_v200.f*. It uses many auxiliary files. All these files must be in a directory called “CODE”, found in the directory where you work, called “F02” for example. The table A.1 briefly describes the parameters files :

survey_limits.h	Definition of the redshift slice Update : see part A.4.3
datafilename.h	Definition of the input files + list of pointings + random catalogues Update : see part A.4.3
cosmo.h	Definition of the cosmological parameters Ω_m and Ω_Λ . For GalICS, $\Omega_m = 0.333$ and $\Omega_\Lambda = 0.667$. (part A.4.3)
bin_sizes.h	Definition of the binning (linear and logarithmic)
maskfile.h	Definition of the photometric fits file. Exemple : oldmasks_f02_b4.fits
point	List of the pointings. The coordinates of their center is written in hour,minute,seconde and degree,minute,seconde. The last column is the number of the quadrants used for each pointing.
weight_param.h	Definition of the value theta used to compute the weight
constants.h	Definition of π , c, ...
count_param.h	Definition of the maximal angular size of the field
gauss_shift.h	Parameters to smooth the spectroscopic $N(z)$ to build the random catalogue
randomvectors.h	Size of the quadrants, number of pointings, ...

TAB. A.1 – Parameters files.

For the F02 field, the input files are *f02_ostateczne* (the parent photometric catalogue) and *f02_sspoc* (the spectroscopic catalogue). If you work with the CDFS data, their name must be *cdfs_sspoc* and *cdfs_ostateczne*.

The files *bin_sizes.h*, *maskfile.h* and *point* must be changed/checked by the user. Exemples of files *point* are given by default in the release : *point_for_f02_galics* and *point_for_f02_vvds*. The 5 last files of table A.1 should not be changed.

In case of GalICS, the parent photometric and spectroscopic files are stored in GALICS/INPUT_FILES (see part A.5). Whatever GalICS or VVDS data, the format of these files are :

*_sspoc	ID	α	δ	redshift	I_{AB}
*_ostateczne	ID	α	δ		I_{AB}

The right ascension α and the declination δ are in degree. A directory called

temporary_random

must always be found in the main directory of the code.

A.4.2 To compute the reference $\xi(s)$ and $\xi(r_p, \pi)$

The main program is *xi_cosmo_naked.f*. It uses many auxiliary files. This code is used to compute the reference $\xi(s)$ and $\xi(r_p, \pi)$ from the 50 GALICS cones. This code is in the directory GALICS/CODE_TRUE. Most of these files are the same than those used with the “sspoc” code (part A.4.1). The file *survey_limits.h* stores also the limits in right ascension/declination of the smaller rectangular area that cover all the field of the spectroscopic data.

This time the input file, in GALICS/INPUT_FILES (see part A.5), are, for exemple *cone001_without_sspoc*. Their format is

cone0 ??_without_sspoc	ID	α	δ	redshift	I_{AB}
------------------------	----	----------	----------	----------	----------

A directory called *temporary_random* must always be found in the main directory of the code.

A.4.3 To update files automatically

When you want to use a code 100 times, things must be done automatically if you do not want to change by hand all the parameter files. For a given redshift bin, you must change the file *datafilename.h* that contains the name of the input files. You must also prepare the different input bootstrap files. Everything is done through the macros (part A.2). They used theses following small programs to update the files for you.

the survey limits

When you run a macro in a redshift bin $[z_1 - z_2]$, the program *modif_range.f* is used to update the file *survey_limits.h*. The name of the original file is changed in *survey_limits_old.h*.

the name of the input files

Whatever you compute $\xi(s)$ with the VVDS data in only one redshift bin or you decide to do 100 bootstrap loops on you data, you must update the name of the input file in *datafilename.h*. It is done with a call to *modif_filename.f*. It uses the file *tail_datafilename.h*. The old file is stored in *datafilename_old.h*.

the cosmology

The values of Ω_m and Ω_Λ are 2 parameters of macros. The small macro *macro_cosmo*, found in the directory of the main code, updates the parameter file *cosmo.h*. In case of GALICS, the user can not change the cosmology ($\Omega_m = 0.333$ and $\Omega_\Lambda = 0.667$).

bootstrap input files

The following files are used to generate a new input file, using a bootstrap resampling method :

- GenRandom.pl (program in Perl)
- make_input_file.f90 (main program)
- param.h (definition of the z-bin, the number of galaxies, ...)

There are N_G galaxies of the original spectroscopic catalogue in the considered redshift bin. The code takes randomly N_G galaxies in this list of N_G galaxies to build a new input file. Several galaxies are chosen twice or more whereas some others are not selected.

A.4.4 To fit $w_p(r_p)$

As briefly explained in part A.3.1, the code *fitwplog_v7.f* computes the best parameters for the fit of $w_p(r_p)$, computed from a given $\xi(r_p, \pi)$ (your data).

The fit is done with a classical and a generalised χ^2 test, using a set of $w_p(r_p)$ (bootstrap resampling of the original data or the 50 GallCS cones). This set of functions is used to compute the errors on the parameters of the fit (correlation length r_0 , slope γ), taking into account the dispersion of $w_p(r_p)$ in each r_p bin, or the complex covariance matrix.

A.4.5 To measure the 1σ errors for r_0 and γ

The program *error_v3.f* is called in *macro_fit_v3*. It computes the “projected” 1σ error of the 2 parameters r_0 and γ . The result is written in an ASCII file, *parameter...dat*. The name of this file depends if the computation is made with bootstrap, GallCS, ...

The first error contour in the plan (r_0, γ) is the 68.3% confidence level, a 1σ limit with 2 free parameters. In order to compute the 1σ error for only one of these parameters, let's say r_0 , the code searches the limits of the projection of the $\chi^2_{\min} + 1$ confidence level contour on the r_0 axis. These values correspond to $r_0 + 1\sigma$ and $r_0 - 1\sigma$. The same exercise is done for γ .

A.4.6 To smooth $\xi(r_p, \pi)$

In the directory MACRO_FIT, the macro *macro_smoothing_xirppi* allows to smooth $\xi(r_p, \pi)$ with an isotropic filter of width $3 h^{-1} \text{Mpc}$. It takes 4 parameters :

macro_smoothing_xirppi [z₁] [z₂] [flag] [Field]

With

1. z₁, minimum redshift of the bin
2. z₂, maximum redshift of the bin
3. – 1 → smooth $\xi(r_p, \pi)$ of bootstrap computation
 - 2 → smooth $\xi(r_p, \pi)$ of GallCS sspoc
 - 3 → smooth $\xi(r_p, \pi)$ of GallCS true
 - 4 → smooth $\xi(r_p, \pi)$ of VVDS data
4. field (F02, CDFS, ...)

The program *scheetman.f* is called and smooth the linear-linear $\xi(r_p, \pi)$. The output is stored in the same directory than the input $\xi(r_p, \pi)$, with the same name but ended with “.smoothed”.

A.4.7 To compute $w_p(r_p)$

In order to compute and store $w_p(r_p)$ in an ascii file, the reader can use the macro *macro_compute_wp*.

Up to now, this macro allows only to compute $w_p(r_p)$ in case of GalICS “true” and GalICS “sspoc”. The syntaxe and the 4 parameters of the macro are :

```
macro_compute_wp [z1] [z2] [Field] [flag]
```

With

1. z_1 , minimum redshift of the bin
2. z_2 , maximum redshift of the bin
3. field (F02, CDFS, ...)
4. – 1 \rightarrow compute $w_p(r_p)$ for $\xi(r_p, \pi)$ of GalICS true
 – 2 \rightarrow compute $w_p(r_p)$ for $\xi(r_p, \pi)$ of GalICS sspoc

The program calls *compute_wp.f* and the output files are stored in the same directory than the input $\xi(r_p, \pi)$, with a name “wp_conexxx.dat”.

A.5 GalICS input files

For each configuration, with GalICS, you need 3 types of files, explained below. An exemple of these files can be found on the same web page where you found this documentation. These files store the GalICS input files with the VVDS-F02 field configuration. For this exemple, the name of the files are :

- *galics_f02_without_sspoc_nov2004.tar.gz*
- *galics_f02_sspoc_nov2004.tar.gz*
- *galics_f02_ostateczne_nov2004.tar.gz*

These 3 files contains the 50 GalICS cones input files for the VVDS-F02 field geometry. They must be put in the directory GALICS. There are 3 ASCII files per cone, as you can check when you untar them by

```
tar -zxvf [file].tar.gz
```

Fifty directories called INPUT_FILES/conexxxx (with xxx=001, 002, ..., 050) will be created.

The files are stored in INPUT_FILES/conexxxx/F02/. In case of a CDFS geometry (exemple), files must be put in INPUT_FILES/conexxxx/CDFS, with the same name as F02’s files (with *f02* changed in *cdfs*).

The redshifts written in the different files are the apparent redshifts found in the GalICS database⁵ + a random velocity dispersion on the line of sight equal to 300 km.s⁻¹ to simulate the VIMOS resolution.

⁵see <http://galics.cosmologie.fr>

A.5.1 Spectroscopic catalogues

For each GalICS cone, there is a final spectroscopic catalogue which is the result of photometric mask, sspoc selection, magnitude incompleteness (as a function of magnitude I_{AB} ; it corresponds to flags 2,3,4 and 9).

The sspoc selection of the files found in *galics_f02_sspoc_nov2004.tar.gz* is the same than for the VVDS-F02 first epoch data, with 20 pointings. The only difference is that all the 80 quadrants are considered as valid.

$$17.5 \leq I_{AB} \leq 24$$

cone0 ??_galics_sspoc	ID	α	δ	redshift	I_{AB}
-----------------------	----	----------	----------	----------	----------

A.5.2 Parent photometric catalogues

The photometric catalogue, with photometric mask + stars (ID>900000)

$$17.5 \leq I_{AB} \leq 24$$

cone ???_ostateczne	ID	α	δ	I_{AB}
---------------------	----	----------	----------	----------

The stars come from simulation of [Robin et al. \(2003\)](#). See the technical paper [Pollo et al. \(2005\)](#) for more information.

A.5.3 One square degree catalogues

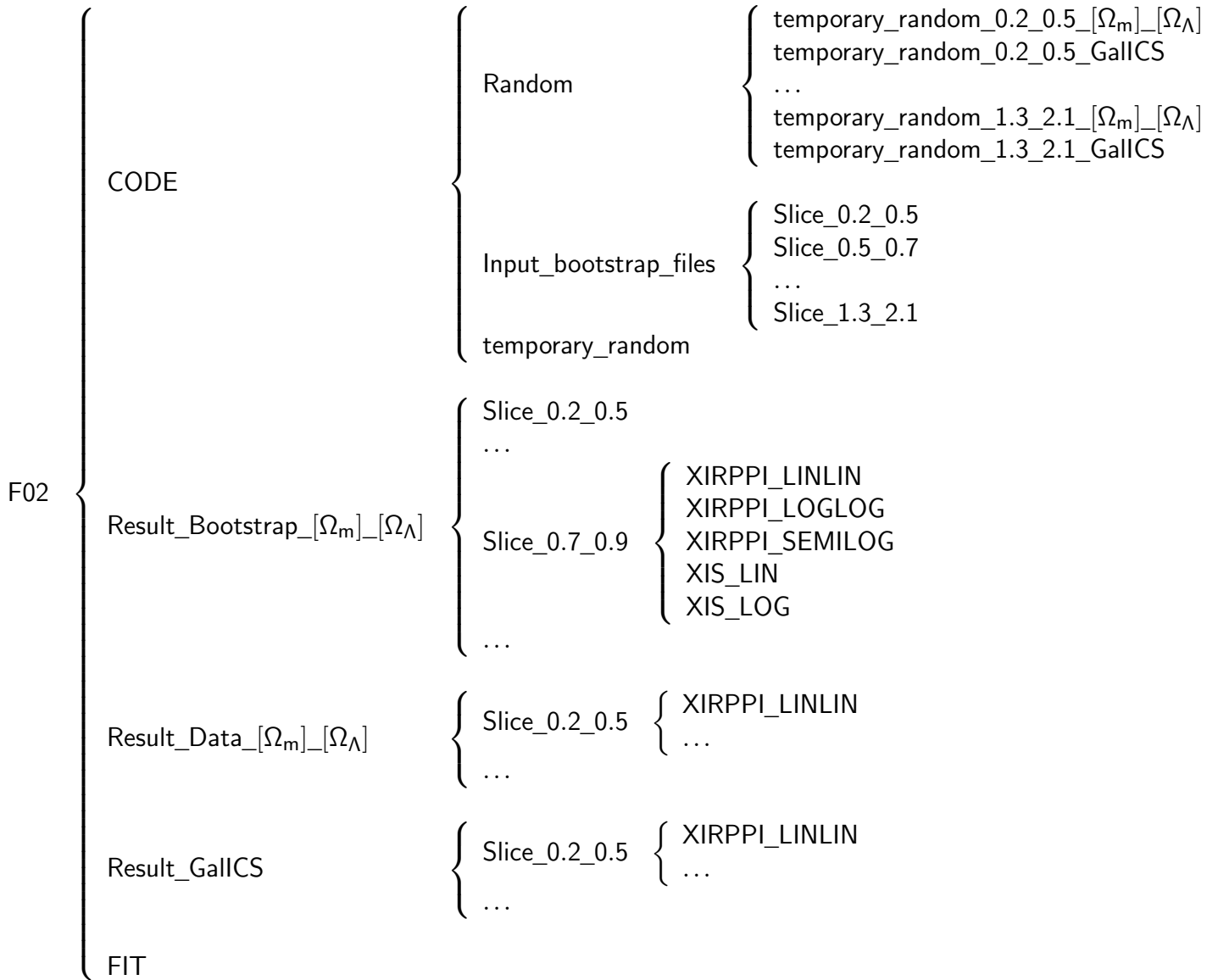
The all 1 by 1 degree field of each GalICS cone + stars (ID>900000). The coordinates are overlapping the F02 field.

$$17.5 \leq I_{AB} \leq 24$$

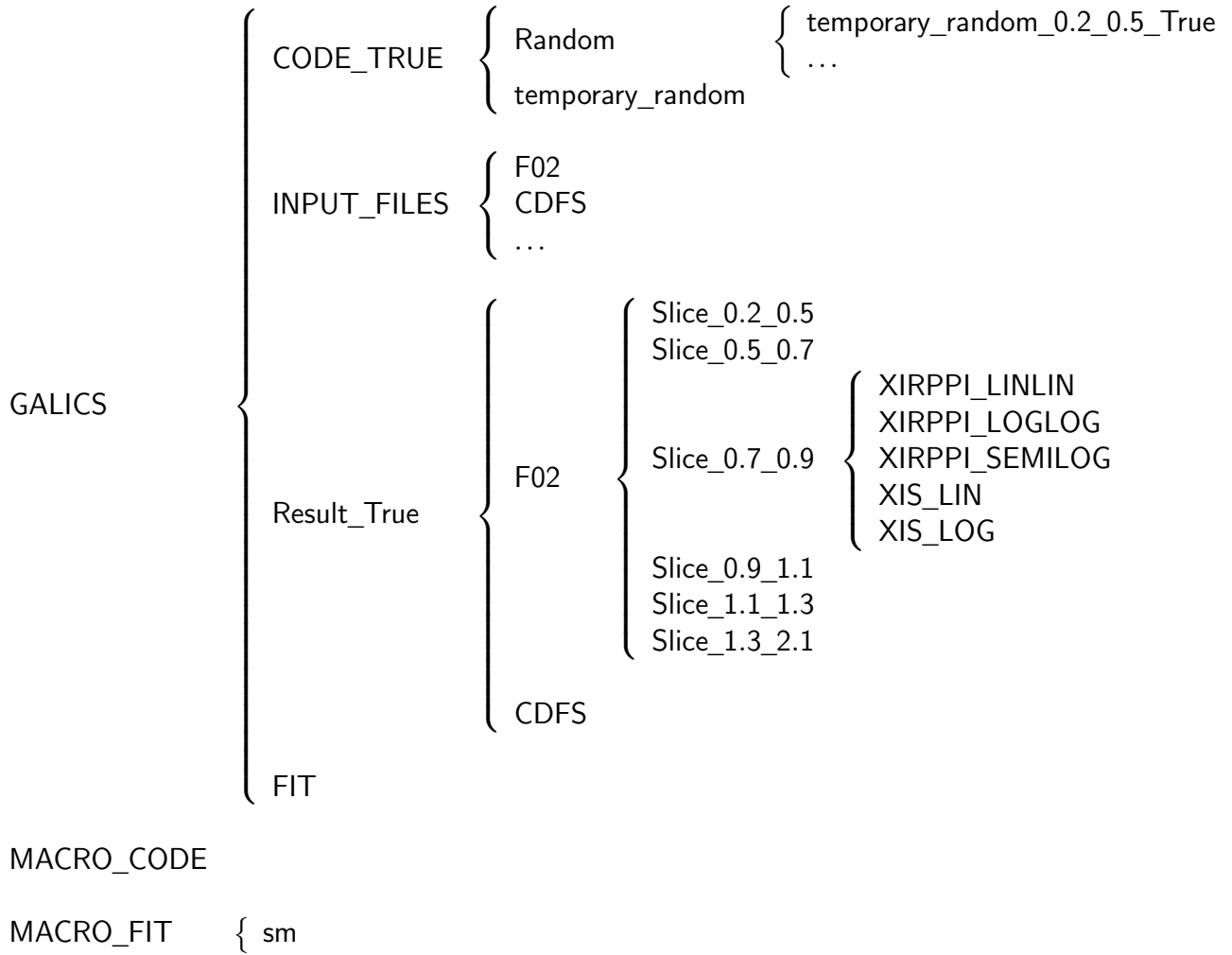
cone0 ??_without_sspoc	ID	α	δ	redshift	I_{AB}
------------------------	----	----------	----------	----------	----------

A.6 Tree structure

A.6.1 directory of your data



A.6.2 directories of the GALICS data and the macros



Bibliographie

- Abazajian, K., Adelman-McCarthy, J. K., Agüeros, M. A., Allam, S. S., *et al.*, 2004, *AJ*, **128**, 502, [ADS](#). [11](#)
- Abell, G. O., 1958, *ApJS*, **3**, 211, [ADS](#). [27](#)
- Adami, C., Mazure, A., Ilbert, O., Cappi, A., *et al.*, 2005, *A&A*, **443**, 805, [ADS](#). [77](#)
- Adelberger, K. L., Steidel, C. C., Giavalisco, M., Dickinson, M., Pettini, M., *et al.*, 1998, *ApJ*, **505**, 18, [ADS](#). [14](#), [15](#), [41](#)
- Adelberger, K. L., Steidel, C. C., Pettini, M., Shapley, A. E., Reddy, N. A., *et al.*, 2005, *ApJ*, **619**, 697, [ADS](#). [41](#), [89](#)
- Adelman-McCarthy, J. K., 2005, *ArXiv Astrophysics e-prints*, [astro-ph/0507711](#). [50](#)
- Allen, P. D., Moustakas, L. A., Dalton, G., MacDonald, E., *et al.*, 2005, *MNRAS*, **360**, 1244, [ADS](#). [41](#)
- Arnouts, S., Cristiani, S., Moscardini, L., Matarrese, S., *et al.*, 1999, *MNRAS*, **310**, 540, [ADS](#). [10](#), [39](#), [40](#)
- Arnouts, S., Moscardini, L., Vanzella, E., Colombi, S., *et al.*, 2002, *MNRAS*, **329**, 355, [ADS](#). [39](#)
- Arnouts, S., Schiminovich, D., Ilbert, O., Tresse, L., *et al.*, 2005, *ApJ*, **619**, L43, [ADS](#). [162](#)
- Arnouts, S., Vandame, B., Benoist, C., Groenewegen, M. A. T., *et al.*, 2001, *A&A*, **379**, 740, [ADS](#). [51](#), [52](#), [76](#)
- Bahcall, N. A., Dong, F., Hao, L., Bode, P., *et al.*, 2003, *ApJ*, **599**, 814, [ADS](#). [5](#), [27](#), [28](#)
- Bahcall, N. A. *et al.*, 1983, *ApJ*, **270**, 20, [ADS](#). [27](#)
- Barrow, J. D., Bhavsar, S. P., *et al.*, 1984, *MNRAS*, **210**, 19P, [ADS](#). [111](#)
- Basilakos, S. *et al.*, 2004, *MNRAS*, **349**, 882, [ADS](#). [27](#)
- Bell, E. F., Wolf, C., Meisenheimer, K., Rix, H.-W., *et al.*, 2004, *ApJ*, **608**, 752. [146](#)
- Benoist, C., Maurogordato, S., da Costa, L. N., Cappi, A., *et al.*, 1996, *ApJ*, **472**, 452, [ADS](#). [28](#), [30](#)
- Benson, A. J., Frenk, C. S., Baugh, C. M., Cole, S., *et al.*, 2001, *MNRAS*, **327**, 1041, [ADS](#). [29](#), [31](#), [44](#), [47](#), [130](#)
- Berlind, A. A. *et al.*, 2002, *ApJ*, **575**, 587, [ADS](#). [47](#)
- Berlind, A. A., Weinberg, D. H., Benson, A. J., Baugh, C. M., *et al.*, 2003, *ApJ*, **593**, 1, [ADS](#). [47](#)
- Blaizot, J., Guiderdoni, B., Devriendt, J. E. G., Bouchet, F. R., Hatton, S. J., *et al.*, 2004, *MNRAS*, **352**, 571, [ADS](#). [101](#)
- Blaizot, J., Wadadekar, Y., Guiderdoni, B., Colombi, S. T., *et al.*, 2005, *MNRAS*, **360**, 159, [ADS](#). [101](#), [102](#)
- Blanton, M., Cen, R., Ostriker, J. P., Strauss, M. A., *et al.*, 2000, *ApJ*, **531**, 1, [ADS](#). [47](#), [169](#)
- Borgani, S. *et al.*, 2001, *Nature*, **409**, 39, [ADS](#). [27](#)

- Bottini, D., Garilli, B., Maccagni, D., Tresse, L., *et al.*, 2005, *PASP*, **117**, 996, [ADS](#). 57, 103
- Brinchmann, J., Abraham, R., Schade, D., Tresse, L., *et al.*, 1998, *ApJ*, **499**, 112, [ADS](#). 20
- Brown, M. J. I., Dey, A., Jannuzi, B. T., Lauer, T. R., Tiede, G. P., et Mikles, V. J., 2003, *ApJ*, **597**, 225, [ADS](#). 39
- Brown, M. J. I., Webster, R. L., et Boyle, B. J., 2000, *MNRAS*, **317**, 782, [ADS](#). 20, 33, 146
- Bruzual, A. G. et Charlot, S., 1993, *ApJ*, **405**, 538, [ADS](#). 140
- Budavári, T., Connolly, A. J., Szalay, A. S., Szapudi, I., *et al.*, 2003, *ApJ*, **595**, 59, [ADS](#). 21, 28, 30, 36, 140
- Budavári, T., Csabai, I., Szalay, A. S., Connolly, A. J., *et al.*, 2001, *AJ*, **122**, 1163, [ADS](#). 62
- Carlberg, R. G., Cowie, L. L., Songaila, A., et Hu, E. M., 1997, *ApJ*, **484**, 538, [ADS](#). 42
- Carlberg, R. G., Yee, H. K. C., Morris, S. L., Lin, H., *et al.*, 2000, *ApJ*, **542**, 57, [ADS](#). 39
- Cimatti, A., Daddi, E., Mignoli, M., Pozzetti, L., *et al.*, 2002, *A&A*, **381**, L68, [ADS](#). 23, 76
- Coil, A. L., Davis, M., Madgwick, D. S., Newman, J. A., *et al.*, 2004, *ApJ*, **609**, 525, [ADS](#). 50, 128, 130, 199
- Coil, A. L., Davis, M., et Szapudi, I., 2001, *PASP*, **113**, 1312, [ADS](#). 100
- Colberg, J. M., White, S. D. M., Yoshida, N., MacFarland, T. J., *et al.*, 2000, *MNRAS*, **319**, 209, [ADS](#). 47, 98
- Cole, S., Hatton, S., Weinberg, D. H., et Frenk, C. S., 1998, *MNRAS*, **300**, 945, [ADS](#). 43, 100
- Coleman, G. D., Wu, C.-C., et Weedman, D. W., 1980, *ApJS*, **43**, 393, [ADS](#). 21, 22, 62, 140
- Colless, M., Dalton, G., Maddox, S., Sutherland, W., *et al.*, 2001, *MNRAS*, **328**, 1039, [ADS](#). 11, 12, 13, 50, 51
- Conti, G., Mattaini, E., Chiappetti, L., Maccagni, D., *et al.*, 2001, *PASP*, **113**, 452, [ADS](#). 55
- da Costa, L. N., Pellegrini, P. S., Davis, M., Meiksin, A., Sargent, W. L. W., et Tonry, J. L., 1991, *ApJS*, **75**, 935, [ADS](#). 10
- Daddi, E., Broadhurst, T., Zamorani, G., Cimatti, A., Röttgering, H., et Renzini, A., 2001, *A&A*, **376**, 825, [ADS](#). 42, 43
- Daddi, E., Cimatti, A., Pozzetti, L., Hoekstra, H., *et al.*, 2000, *A&A*, **361**, 535, [ADS](#). 42
- Dalton, G. B., Croft, R. A. C., Efstathiou, G., Sutherland, W. J., Maddox, S. J., et Davis, M., 1994, *MNRAS*, **271**, L47+, [ADS](#). 27
- Davis, M., Efstathiou, G., Frenk, C. S., et White, S. D. M., 1985, *ApJ*, **292**, 371, [ADS](#). 24
- Davis, M., Faber, S. M., Newman, J., Phillips, A. C., *et al.*, dans *Discoveries and Research Prospects from 6- to 10-Meter-Class Telescopes II. Edited by Guhathakurta, Puragra. Proceedings of the SPIE, Volume 4834, pp. 161-172 (2003).* (2003), 161–172, [ADS](#). 10, 49, 50, 53
- Davis, M. et Geller, M. J., 1976, *ApJ*, **208**, 13, [ADS](#). 18, 31
- Davis, M. et Peebles, P. J. E., 1983, *ApJ*, **267**, 465, [ADS](#). 16, 26
- de Lapparent, V., Geller, M. J., et Huchra, J. P., 1986, *ApJ*, **302**, L1, [ADS](#). 10, 11
- de Vaucouleurs, G., dans *IAU Symp. 58 : The Formation and Dynamics of Galaxies* (1974), 335–+, [ADS](#). 18
- Dressler, A., 1980, *ApJ*, **236**, 351, [ADS](#). 31, 32, 44
- Eke, V. R., Cole, S., et Frenk, C. S., 1996, *MNRAS*, **282**, 263, [ADS](#). 100
- Evrard, A. E., MacFarland, T. J., Couchman, H. M. P., Colberg, J. M., *et al.*, 2002, *ApJ*, **573**, 7, [ADS](#). 98
- Firth, A. E., Somerville, R. S., McMahon, R. G., Lahav, O., *et al.*, 2002, *MNRAS*, **332**, 617, [ADS](#). 42, 140

- Fisher, K. B., Davis, M., Strauss, M. A., Yahil, A., et Huchra, J. P., 1994, *MNRAS*, **267**, 927, [ADS](#). [26](#), [81](#), [82](#)
- Fisher, K. B., Huchra, J. P., Strauss, M. A., Davis, M., Yahil, A., et Schlegel, D., 1995, *ApJS*, **100**, 69, [ADS](#). [11](#)
- Foucaud, S., McCracken, H. J., Le Fèvre, O., Arnouts, S., *et al.*, 2003, *A&A*, **409**, 835, [ADS](#). [41](#)
- Frenk, C. S., Colberg, J. M., H. M. P., Couchman, Efstathiou, *et al.*, 2000, *ArXiv Astrophysics e-prints*, [astro-ph/0007362](#). [98](#)
- Fry, J. N., 1996, *ApJ*, **461**, L65+, [ADS](#). [45](#), [169](#)
- Garilli, B. et l'équipe VVDS, 2005, *A&A*, **000**, 000. [79](#)
- Gaztañaga, E. et Juszkiewicz, R., 2001, *ApJ*, **558**, L1, [ADS](#). [16](#)
- Giacconi, R., Zirm, A., Wang, J., Rosati, P., *et al.*, 2002, *ApJS*, **139**, 369, [ADS](#). [75](#)
- Giallongo, E., Salimbeni, S., Menci, N., Zamorani, G., *et al.*, 2005, *ApJ*, **622**, 116, [ADS](#). [146](#), [147](#)
- Gialvalisco, M. et Dickinson, M., 2001, *ApJ*, **550**, 177, [ADS](#). [41](#)
- Gialvalisco, M., Ferguson, H. C., Koekemoer, A. M., Dickinson, M., *et al.*, 2004, *ApJ*, **600**, L93, [ADS](#). [10](#), [50](#), [76](#)
- Gialvalisco, M., Steidel, C. C., Adelberger, K. L., Dickinson, M. E., Pettini, M., et Kellogg, M., 1998, *ApJ*, **503**, 543, [ADS](#). [41](#)
- Giovanelli, R. et Haynes, M. P., 1985, *AJ*, **90**, 2445, [ADS](#). [10](#)
- Giovanelli, R., Haynes, M. P., et Chincarini, G. L., 1986, *ApJ*, **300**, 77, [ADS](#). [31](#)
- Gray, A. G., Moore, A. W., Nichol, R. C., Connolly, A. J., Genovese, C., et Wasserman, L., dans *ASP Conf. Ser. 314 : Astronomical Data Analysis Software and Systems (ADASS) XIII* (2004), 249–+, [ADS](#). [17](#)
- Groth, E. J. et Peebles, P. J. E., 1977, *ApJ*, **217**, 385, [ADS](#). [16](#)
- Guzzo, L., Bartlett, J. G., Cappi, A., Maurogordato, S., *et al.*, 2000, *A&A*, **355**, 1, [ADS](#). [26](#), [28](#), [30](#)
- Guzzo, L., Iovino, A., Chincarini, G., Giovanelli, R., et Haynes, M. P., 1991, *ApJ*, **382**, L5, [ADS](#). [16](#)
- Guzzo, L., Strauss, M. A., Fisher, K. B., Giovanelli, R., et Haynes, M. P., 1997, *ApJ*, **489**, 37, [ADS](#). [28](#), [30](#), [32](#), [33](#), [42](#)
- Hamilton, A. J. S., 1988, *ApJ*, **331**, L59, [ADS](#). [28](#)
- Hamilton, A. J. S., 1993, *ApJ*, **417**, 19, [ADS](#). [17](#)
- Hatton, S., Devriendt, J. E. G., Ninin, S., Bouchet, F. R., Guiderdoni, B., et Vibert, D., 2003, *MNRAS*, **343**, 75, [ADS](#). [100](#), [101](#)
- Hawkins, E., Maddox, S., Cole, S., Lahav, O., *et al.*, 2003, *MNRAS*, **346**, 78, [ADS](#). [26](#), [27](#)
- Hermit, S., Santiago, B. X., Lahav, O., Strauss, M. A., *et al.*, 1996, *MNRAS*, **283**, 709, [ADS](#). [32](#), [33](#)
- Hernquist, L., 1990, *ApJ*, **356**, 359, [ADS](#). [104](#)
- Hubble, E. P., 1936, *Yale University Press*, [ADS](#). [18](#), [31](#)
- Huchra, J., Davis, M., Latham, D., et Tonry, J., 1983, *ApJS*, **52**, 89, [ADS](#). [10](#)
- Ilbert, O., 2004, *Formation et évolution de populations de galaxies à grand décalage spectral - Fonctions de luminosité du relevé VVDS*, thèse de doctorat, Université de Provence - Aix-Marseille I, [ADS](#). [140](#)
- Iovino, A., Giovanelli, R., Haynes, M., Chincarini, G., et Guzzo, L., 1993, *MNRAS*, **265**, 21, [ADS](#). [32](#)

- Iovino, A., McCracken, H. J., Garilli, B., Foucaud, S., *et al.*, 2005, *A&A*, **442**, 423, [ADS](#). 51, 52
- Jenkins, A., Frenk, C. S., Pearce, F. R., Thomas, P. A., *et al.*, 1998, *ApJ*, **499**, 20, [ADS](#). 43
- Jing, Y. P., Mo, H. J., et Boerner, G., 1998, *ApJ*, **494**, 1, [ADS](#). 26
- Kaiser, N., 1984, *ApJ*, **284**, L9, [ADS](#). 7, 24
- Kaiser, N., 1987, *MNRAS*, **227**, 1, [ADS](#). 81, 82
- Kauffmann, G., Colberg, J. M., Diaferio, A., et White, S. D. M., 1999a, *MNRAS*, **303**, 188, [ADS](#). 44, 47, 99, 208
- Kauffmann, G., Colberg, J. M., Diaferio, A., et White, S. D. M., 1999b, *MNRAS*, **307**, 529, [ADS](#). 44, 45
- Kauffmann, G., Nusser, A., et Steinmetz, M., 1997, *MNRAS*, **286**, 795, [ADS](#). 43
- Kennicutt, R. C., 1992, *ApJS*, **79**, 255, [ADS](#). 62
- Landy, S. D. et Szalay, A. S., 1993, *ApJ*, **412**, 64, [ADS](#). 17
- Landy, S. D., Szalay, A. S., et Koo, D. C., 1996, *ApJ*, **460**, 94, [ADS](#). 20, 33
- Lawrence, A., Rowan-Robinson, M., Ellis, R. S., Frenk, C. S., *et al.*, 1999, *MNRAS*, **308**, 897, [ADS](#). 11
- Le Fèvre, O., Crampton, D., Felenbok, P., et Monnet, G., 1994a, *A&A*, **282**, 325, [ADS](#). 9
- Le Fèvre, O., Crampton, D., Hammer, F., Lilly, S. J., et Tresse, L., 1994b, *ApJ*, **423**, L89+, [ADS](#). 39
- Le Fèvre, O., Crampton, D., Lilly, S. J., Hammer, F., et Tresse, L., 1995, *ApJ*, **455**, 60, [ADS](#). 9, 55, 66
- Le Fèvre, O., Guzzo, L., Meneux, B., Pollo, A., *et al.*, 2005a, *A&A*, **439**, 877, [ADS](#). 129
- Le Fèvre, O., Hudon, D., Lilly, S. J., Crampton, D., Hammer, F., et Tresse, L., 1996, *ApJ*, **461**, 534, [ADS](#). 14, 39, 40
- Le Fèvre, O., Mellier, Y., McCracken, H. J., Foucaud, S., *et al.*, 2004a, *A&A*, **417**, 839, [ADS](#). 51, 53
- Le Fèvre, O., Paltani, S., Arnouts, S., Charlot, S., *et al.*, 2005b, *Nature*, **437**, 519, [ADS](#). 210
- Le Fèvre, O., Saisse, M., Mancini, D., Brau-Nogue, S., *et al.*, dans *Instrument Design and Performance for Optical/Infrared Ground-based Telescopes. Edited by Iye, Masanori; Moorwood, Alan F. M. Proceedings of the SPIE, Volume 4841, pp. 1670-1681 (2003).* (2003), 1670–1681, [ADS](#). 55
- Le Fèvre, O., Vettolani, G., Garilli, B., Tresse, L., *et al.*, 2005c, *A&A*, **439**, 845, [ADS](#). 10, 73, 89, 92
- Le Fèvre, O., Vettolani, G., Paltani, S., Tresse, L., *et al.*, 2004b, *A&A*, **428**, 1043, [ADS](#). 73, 76
- Lee, K., Giavalisco, M., Gnedin, O. Y., Somerville, R., *et al.*, 2005, *ArXiv Astrophysics e-prints*, [astro-ph/0508090](#). 41
- Lilly, S. J., Le Fèvre, O., Crampton, D., Hammer, F., et Tresse, L., 1995a, *ApJ*, **455**, 50, [ADS](#). 50
- Lilly, S. J., Tresse, L., Hammer, F., Crampton, D., et Le Fèvre, O., 1995b, *ApJ*, **455**, 108, [ADS](#). 9
- Limber, D. N., 1953, *ApJ*, **117**, 134, [ADS](#). 41
- Longair, M. S., rédacteur, *Galaxy formation* (Springer. (Astronomy and astrophysics library) QB981 .L846 1998, 1998), [ADS](#). 92
- Loveday, J., Maddox, S. J., Efstathiou, G., et Peterson, B. A., 1995, *ApJ*, **442**, 457, [ADS](#). 26, 28, 30, 32, 33

- Loveday, J., Peterson, B. A., Maddox, S. J., et Efstathiou, G., 1996, *ApJS*, **107**, 201, [ADS](#).
[11](#)
- Loveday, J., Tresse, L., et Maddox, S., 1999, *MNRAS*, **310**, 281, [ADS](#). [21](#), [36](#), [145](#), [146](#)
- Madgwick, D. S., Hawkins, E., Lahav, O., Maddox, S., *et al.*, 2003, *MNRAS*, **344**, 847, [ADS](#).
[36](#)
- Madgwick, D. S., Lahav, O., Baldry, I. K., Baugh, C. M., *et al.*, 2002, *MNRAS*, **333**, 133, [ADS](#). [21](#)
- Marinoni, C., Le Fèvre, O., Meneux, B., Iovino, A., *et al.*, 2005, *A&A*, **442**, 801, [ADS](#). [51](#),
[168](#), [169](#)
- Martin, D. C., Fanson, J., Schiminovich, D., Morrissey, P., *et al.*, 2005, *ApJ*, **619**, L1, [ADS](#).
[162](#)
- Marzke, R. O., Geller, M. J., da Costa, L. N., et Huchra, J. P., 1995, *AJ*, **110**, 477, [ADS](#). [26](#)
- McCracken, H. J., Le Fèvre, O., Brodwin, M., Foucaud, S., *et al.*, 2001, *A&A*, **376**, 756, [ADS](#). [8](#)
- McCracken, H. J., Radovich, M., Bertin, E., Mellier, Y., *et al.*, 2003, *A&A*, **410**, 17, [ADS](#). [51](#),
[52](#), [103](#)
- Mo, H. J., Jing, Y. P., et Boerner, G., 1992, *ApJ*, **392**, 452, [ADS](#). [111](#)
- Mo, H. J. et White, S. D. M., 1996, *MNRAS*, **282**, 347, [ADS](#). [43](#), [47](#), [169](#)
- Morrissey, P., Schiminovich, D., Barlow, T. A., Martin, D. C., *et al.*, 2005, *ApJ*, **619**, L7, [ADS](#).
[52](#)
- Moscardini, L., Coles, P., Lucchin, F., et Matarrese, S., 1998, *MNRAS*, **299**, 95, [ADS](#). [45](#)
- Neuschaefer, L. W., Ratnatunga, K. U., Griffiths, R. E., Casertano, S., et Im, M., 1995, *ApJ*,
453, 559, [ADS](#). [20](#), [32](#)
- Norberg, P., Baugh, C. M., Hawkins, E., Maddox, S., *et al.*, 2001, *MNRAS*, **328**, 64, [ADS](#).
[29](#), [30](#), [31](#)
- Norberg, P., Baugh, C. M., Hawkins, E., Maddox, S., *et al.*, 2002, *MNRAS*, **332**, 827, [ADS](#).
[13](#), [21](#), [36](#), [37](#), [145](#), [146](#)
- Ouchi, M., Shimasaku, K., Okamura, S., Doi, M., *et al.*, 2001, *ApJ*, **558**, L83, [ADS](#). [41](#)
- Ouchi, M., Shimasaku, K., Okamura, S., Furusawa, H., *et al.*, 2004, *ApJ*, **611**, 685, [ADS](#).
[41](#)
- Peacock, J. A., *Cosmological physics* (Cosmological physics. Publisher : Cambridge, UK :
Cambridge University Press, 1999. ISBN : 0521422701, 1999), [ADS](#). [4](#), [24](#), [92](#)
- Peacock, J. A., Cole, S., Norberg, P., Baugh, C. M., *et al.*, 2001, *Nature*, **410**, 169, [ADS](#). [82](#)
- Peebles, P. J. E., *The large-scale structure of the universe* (Research supported by the
National Science Foundation. Princeton, N.J., Princeton University Press, 1980. 435
p., 1980), [ADS](#). [25](#), [82](#)
- Peng, C. Y., Ho, L. C., Impey, C. D., et Rix, H.-W., 2002, *AJ*, **124**, 266, [ADS](#). [20](#)
- Phleps, S. et Meisenheimer, K., 2003, *A&A*, **407**, 855, [ADS](#). [10](#)
- Phleps, S., Peacock, J. A., Meisenheimer, K., et Wolf, C., 2005, *ArXiv Astrophysics e-prints*, [astro-ph/0506320](#). [16](#), [145](#)
- Pickles, A. J., 1985, *ApJS*, **59**, 33, [ADS](#). [65](#)
- Pollo, A., Meneux, B., Guzzo, L., Le Fèvre, O., *et al.*, 2005, *A&A*, **439**, 887, [ADS](#). [92](#), [109](#),
[111](#), [221](#)
- Postman, M. et Geller, M. J., 1984, *ApJ*, **281**, 95, [ADS](#). [32](#)
- Postman, M., Lauer, T. R., Szapudi, I., et Oegerle, W., 1998, *ApJ*, **506**, 33, [ADS](#). [39](#)
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., et Flannery, B. P., *Numerical recipes in*

- FORTTRAN. The art of scientific computing* (Cambridge : University Press, 1992, 2nd ed., 1992), [ADS](#). 94
- Radovich, M., Arnaboldi, M., Ripepi, V., Massarotti, M., *et al.*, 2004, *A&A*, **417**, 51, [ADS](#). 51, 52
- Ratcliffe, A., Shanks, T., Parker, Q. A., Broadbent, A., *et al.*, 1998a, *MNRAS*, **300**, 417, [ADS](#). 11
- Ratcliffe, A., Shanks, T., Parker, Q. A., et Fong, R., 1998b, *MNRAS*, **296**, 173, [ADS](#). 26
- Rix, H.-W., Barden, M., Beckwith, S. V. W., Bell, E. F., *et al.*, 2004, *ApJS*, **152**, 163, [ADS](#). 50
- Robin, A. C., Reyl  , C., Derri  re, S., et Picaud, S., 2003, *A&A*, **409**, 523, [ADS](#). 103, 104, 221
- Roche, N., Shanks, T., Metcalfe, N., et Fong, R., 1996, *MNRAS*, **280**, 397, [ADS](#). 20, 33
- Santiago, B. X. et Strauss, M. A., 1992, *ApJ*, **387**, 9, [ADS](#). 32
- Santiago, B. X., Strauss, M. A., Lahav, O., Davis, M., Dressler, A., et Huchra, J. P., 1995, *ApJ*, **446**, 457, [ADS](#). 10
- Saunders, W., Rowan-Robinson, M., et Lawrence, A., 1992, *MNRAS*, **258**, 134, [ADS](#). 26
- Saunders, W., Sutherland, W. J., Maddox, S. J., Keeble, O., *et al.*, 2000, *MNRAS*, **317**, 55, [ADS](#). 11
- Scaramella, R. et l'  quipe VVDS, 2005, *A&A*, **000**, 000. 65
- Scodeggio, M., Franzetti, P., Garilli, B., Zanichelli, A., *et al.*, 2005, *PASP*, **117**, 1284, [ADS](#). 58
- Shapley, A. E., Steidel, C. C., Pettini, M., et Adelberger, K. L., 2003, *ApJ*, **588**, 65, [ADS](#). 62
- Shectman, S. A., Landy, S. D., Oemler, A., Tucker, D. L., *et al.*, 1996, *ApJ*, **470**, 172, [ADS](#). 11
- Shepherd, C. W., Carlberg, R. G., Yee, H. K. C., Morris, S. L., *et al.*, 2001, *ApJ*, **560**, 72, [ADS](#). 28, 30, 39, 140
- Simard, L., Willmer, C. N. A., Vogt, N. P., Sarajedini, V. L., *et al.*, 2002, *ApJS*, **142**, 1, [ADS](#). 20
- Somerville, R. S., Lee, K., Ferguson, H. C., Gardner, J. P., Moustakas, L. A., et Giavalisco, M., 2004, *ApJ*, **600**, L171, [ADS](#). 17
- Somerville, R. S., Lemson, G., Sigad, Y., Dekel, A., Kauffmann, G., et White, S. D. M., 2001, *MNRAS*, **320**, 289, [ADS](#). 47, 169
- Somerville, R. S. et Primack, J. R., 1999, *MNRAS*, **310**, 1087, [ADS](#). 47
- Spergel, D. N., Verde, L., Peiris, H. V., Komatsu, E., *et al.*, 2003, *ApJS*, **148**, 175, [ADS](#). 5, 6
- Springel, V., White, S. D. M., Jenkins, A., Frenk, C. S., *et al.*, 2005, *Nature*, **435**, 629, [ADS](#). 43, 98
- Steidel, C. C., Adelberger, K. L., Dickinson, M., Giavalisco, M., Pettini, M., et Kellogg, M., 1998, *ApJ*, **492**, 428, [ADS](#). 14, 40
- Steidel, C. C., Adelberger, K. L., Shapley, A. E., Pettini, M., Dickinson, M., et Giavalisco, M., 2003, *ApJ*, **592**, 728, [ADS](#). 23
- Steidel, C. C., Giavalisco, M., Pettini, M., Dickinson, M., et Adelberger, K. L., 1996, *ApJ*, **462**, L17+, [ADS](#). 22, 53, 62
- Strateva, I., Ivezi  ,   ., Knapp, G. R., Narayanan, V. K., *et al.*, 2001, *AJ*, **122**, 1861, [ADS](#). 21, 146
- Strauss, M. A., Huchra, J. P., Davis, M., Yahil, A., Fisher, K. B., et Tonry, J., 1992, *ApJS*, **83**, 29, [ADS](#). 11

- Tegmark, M. et Peebles, P. J. E., 1998, *ApJ*, **500**, L79+, [ADS](#). 45
- Totsuji, H. et Kihara, T., 1969, *PASJ*, **21**, 221, [ADS](#). 15
- Tucker, D. L., Oemler, A., Kirshner, R. P., Lin, H., *et al.*, 1997, *MNRAS*, **285**, L5, [ADS](#). 26
- van Dokkum, P. G., Franx, M., Fabricant, D., Illingworth, G. D., et Kelson, D. D., 2000, *ApJ*, **541**, 95, [ADS](#). 146
- Vanden Berk, D. E., Richards, G. T., Bauer, A., Strauss, M. A., *et al.*, 2001, *AJ*, **122**, 549, [ADS](#). 62
- Vettolani, G., Zucca, E., Merighi, R., Mignoli, M., *et al.*, 1998, *A&AS*, **130**, 323, [ADS](#). 11
- Weinberg, D. H., Davé, R., Katz, N., et Hernquist, L., 2004, *ApJ*, **601**, 1, [ADS](#). 44, 45, 46, 130, 233
- White, S. D. M., Tully, R. B., et Davis, M., 1988, *ApJ*, **333**, L45, [ADS](#). 32
- Willmer, C. N. A., da Costa, L. N., et Pellegrini, P. S., 1998, *AJ*, **115**, 869, [ADS](#). 20, 26, 28, 30, 32, 33, 35, 42
- Willmer, C. N. A., Faber, S. M., Koo, D. C., Weiner, B. J., *et al.*, 2005, *ArXiv Astrophysics e-prints*, [astro-ph/0506041](#). 146
- Wolf, C., Meisenheimer, K., Kleinheinrich, M., Borch, A., *et al.*, 2004, *A&A*, **421**, 913, [ADS](#). 10, 50
- Xu, C. K., Donas, J., Arnouts, S., Wyder, T. K., *et al.*, 2005, *ApJ*, **619**, L11, [ADS](#). 162
- Yan, R., White, M., et Coil, A. L., 2004, *ApJ*, **607**, 739, [ADS](#). 100
- Yee, H. K. C., Morris, S. L., Lin, H., Carlberg, R. G., *et al.*, 2000, *ApJS*, **129**, 475, [ADS](#). 10, 50
- Yoshida, N., Colberg, J., White, S. D. M., Evrard, A. E., *et al.*, 2001, *MNRAS*, **325**, 803, [ADS](#). 43
- Zehavi, I., Blanton, M. R., Frieman, J. A., Weinberg, D. H., *et al.*, 2002, *ApJ*, **571**, 172, [ADS](#). 20, 21, 26, 28, 29, 30, 33, 34, 35, 36, 145, 146
- Zehavi, I., Eisenstein, D. J., Nichol, R. C., Blanton, M. R., *et al.*, 2005a, *ApJ*, **621**, 22, [ADS](#). 145
- Zehavi, I., Weinberg, D. H., Zheng, Z., Berlind, A. A., *et al.*, 2004, *ApJ*, **608**, 16, [ADS](#). 16, 26
- Zehavi, I., Zheng, Z., Weinberg, D. H., Frieman, J. A., *et al.*, 2005b, *ApJ*, **630**, 1, [ADS](#). 16, 21, 29, 34, 35, 146
- Zhang, L. L. et Pen, U.-L., 2005, *New Astronomy*, **10**, 569, [ADS](#). 17
- Zheng, W., Kriss, G. A., Telfer, R. C., Grimes, J. P., et Davidsen, A. F., 1997, *ApJ*, **475**, 469, [ADS](#). 62
- Zucca, E., Ilbert, O., Bardelli, S., Tresse, L., *et al.*, 2005, *ArXiv Astrophysics e-prints*, [astro-ph/0506393](#). 140, 143, 146, 234

Liste des tableaux

1.1	Longueur de corrélation r_0 en fonction de la luminosité	30
1.2	Longueur de corrélation r_0 en fonction du type morphologique	33
1.3	Longueur de corrélation r_0 en fonction de la couleur	35
1.4	Longueur de corrélation r_0 en fonction du type spectral	36
1.5	Longueur de corrélation r_0 de galaxies <i>Lyman-Break</i>	41
2.1	Paramètres important qui conditionnent l'efficacité d'un sondage	50
2.2	Résumé des données photométriques disponibles pour les différents champs du VVDS	52
2.3	Liste des principales raies en émission et/ou absorption	63
2.4	Définition des notes de confiance attribuées à chaque <i>redshift</i>	70
2.5	Répartition du nombre d'objets en fonction de leur nature et de leur note dans le VVDS-02h	75
2.6	Répartition du nombre d'objets en fonction de leur type et de leur note dans le VVDS-CDFS	77
3.1	$\Delta\chi^2$ en fonction du degré de confiance et du nombre de degrés de liberté ν	95
4.1	Écart relatif entre la longueur de corrélation r_0 (respectivement la pente γ) mesurée et la référence.	110
5.1	Propriété des 4 types de galaxies du champs VVDS-02	141
5.2	Paramètres r_0 et γ mesurés à partir des échantillons GALEX-VVDS sélectionnés en NUV	166
5.3	Paramètres r_0 et γ mesurés à partir des échantillons GALEX-VVDS sélectionnés en FUV	167
6.1	Tableau récapitulatif : mesures obtenues pour la population globale	204
6.2	Tableau récapitulatif : mesures obtenues pour chaque type spectral	205
6.3	Tableau récapitulatif : mesures obtenues pour les populations rouges et bleues	206
6.4	Tableau récapitulatif : mesures obtenues pour chaque sous échantillon VVDS-GALEX	207
A.1	Parameters files.	217

Table des figures

1.1	Relation temps- <i>redshift</i>	5
1.2	“Slice of the Universe”	11
1.3	Projection dans l’espace des galaxies du 2dFGRS	12
1.4	Exemple de structures présentes dans le champ sud du 2dFGRS	13
1.5	Distribution en <i>redshifts</i> du 2dFGRS	13
1.6	Distribution en <i>redshifts</i> des champs du CFRS	14
1.7	Distributions en <i>redshifts</i> de galaxies <i>Lyman-Break</i> à $z \sim 3$	15
1.8	Magnitude absolue en bande B des galaxies du sondage VVDS en fonction du <i>redshift</i>	19
1.9	Classification morphologique de Hubble	19
1.10	Spectres de références CWW	22
1.11	Courbe de transmission des filtres de sélection des galaxies <i>Lyman-Break</i>	23
1.12	Fluctuations d’un champ gaussien aléatoire $\delta(x)$	24
1.13	Longueur de corrélation d’amas de galaxies sélectionnés en optique et en X	27
1.14	Fonction $w_p(r_p)$ du SDSS en fonction de la luminosité	29
1.15	Longueur de corrélation en fonction de la magnitude absolue en bande R	30
1.16	Longueur de corrélation en fonction de la magnitude absolue en bande B	31
1.17	Fonction $w_p(r_p)$ dans le SDSS en fonction de la couleur	34
1.18	Ségrégation des galaxies en fonction de leur couleur	35
1.19	Longueur de corrélation r_0 en fonction de la magnitude absolue b_J pour les deux types de galaxies du 2dFGRS	37
1.20	Longueur de corrélation r_0 en fonction du type de chaque galaxie	38
1.21	Longueur de corrélation r_0 en fonction du <i>redshift</i>	40
1.22	Longueur de corrélation de population de galaxies de type <i>Lyman-Break</i> identifiées à grands <i>redshifts</i> .	42
1.23	Simulation GIF	44
1.24	Évolution de la longueur de corrélation spatiale dans pour 2 modèles cosmologiques	45
1.25	Évolution de la longueur de corrélation spatiale r_0 et pente γ (Weinberg et al., 2004)	46
2.1	Position des quatre champs VVDS sur le ciel	53
2.2	Courbe de transmission des filtres utilisés pour le VVDS	54
2.3	Champ de vue de l’instrument VIMOS	56
2.4	Exemple d’une sélection d’objets par SSPOC	57
2.5	Exemple d’une exposition MOS	58

2.6	Exemple de spectre à 2 dimensions	59
2.7	Profil d'extraction d'un spectre	60
2.8	Visibilité de quelques raies en fonction du <i>redshift</i>	63
2.9	Spectres de référence réalisés à partir des spectres du VVDS	64
2.10	Solutions proposées par KBRED pour un spectre en émission	67
2.11	Solutions proposées par KBRED pour un spectre en absorption	68
2.12	Panneau <i>Slit Summary</i> de VIPGI	69
2.13	Spectres de galaxies du VVDS-02h	71
2.14	Spectres d'étoiles et de quasars du VVDS-02h	72
2.15	Objets observés avec VIMOS dans le VVDS-02h	74
2.16	Distribution en <i>redshift</i> du VVDS-02h	74
2.17	Le taux de succès spectroscopique	75
2.18	Objets observés avec VIMOS dans le CDFS	76
2.19	Distribution en <i>redshift</i> du VVDS-CDFS	77
2.20	Objets observés avec VIMOS dans les champs VVDS-10h et VVDS-14h	78
2.21	Objets observés avec VIMOS dans le champ VVDS-22h	78
2.22	Distributions en <i>redshifts</i> des champs VVDS-10h et VVDS-22h	79
3.1	Fonction $\xi(r_p, \pi)$ du 2dFGRS	82
3.2	Schéma : définition de r_p et π	83
3.3	Masques photométriques	86
3.4	Grille des 5 pointés VIMOS sur le VVDS-CDFS	87
3.5	Schéma : un poids pour retrouver le nombre d'objets	90
3.6	Schéma : un poids pour retrouver le nombre de paires	91
3.7	Schéma : niveaux de confiance	95
4.1	Géométrie angulaire des simulations <i>Hubble-Volume</i> disponibles	99
4.2	Processus de réalisation d'un cône <i>GallCS</i>	102
4.3	Distribution des étoiles dans le VVDS-02h	104
4.4	Distribution moyenne des rayons apparents des objets du VVDS-02h	105
4.5	Pourcentage moyen d'objets sélectionnés par SSPOC	106
4.6	Distribution moyenne en <i>redshift</i> des 50 simulations <i>GallCS</i>	107
4.7	Comptage moyen en bande I des 50 simulations <i>GallCS</i>	107
4.8	$\xi(r_p, \pi)$ mesurée avec et sans correction de la stratégie d'observation	109
4.9	$w_p(r_p)$ mesurée avec et sans correction de la stratégie d'observation	109
4.10	Distribution des différences relatives entre la valeur de r_0 (γ) de référence et celle mesurée dans les conditions du VVDS	110
4.11	Différence relative entre les erreurs sur $w_p(r_p)$ estimées à partir d'un ré-échantillonnage <i>bootstrap</i> des données VVDS et celles calculées à partir des simulations <i>GallCS</i>	112
5.1	Distribution en <i>redshifts</i> du VVDS-02h avec la stratégie DEEP2	128
5.2	Fonction $\xi(r_p, \pi)$ pour l'échantillon spectroscopique global du VVDS-02h	129
5.3	Distributions en <i>redshifts</i> des 4 types de galaxies du VVDS-02h	142
5.4	Distributions en <i>redshifts</i> des 4 types de galaxies définis dans 20 champs <i>GallCS</i> selon la méthode de Zucca et al. (2005)	143

5.5	Distributions en <i>redshifts</i> des types E/S0 et Spirales/Irrégulières définis dans la base de donnée <i>GalICS</i>	144
5.6	Fonctions $\xi(r_p, \pi)$ des échantillons spectroscopiques du VVDS-02h incluant les types 1+2 et 3+4	145
5.7	Division couleur-magnitude dans le champ VVDS-02h	147
5.8	Fonctions $\xi(r_p, \pi)$ des populations rouges et bleues dans le champ VVDS-02h	148
5.9	Fonctions $w_p(r_p)$ des populations rouges et bleues en fonction du <i>redshift</i> .	148
5.10	Comptages en NUV (à gauche) et FUV (à droite)	163
5.11	Diagramme couleur-magnitude en NUV (à gauche) et FUV (à droite)	164
5.12	Distributions en <i>redshift</i> des échantillons GALEX-VVDS	164
5.13	Fonction de corrélation $w_p(r_p)$ mesurée à partir de l'échantillon GALEX-VVDS en NUV	165
5.14	Couleurs en NUV	166
5.15	Fonction de corrélation $w_p(r_p)$ mesurée à partir de l'échantillon GALEX-VVDS en FUV	167
5.16	Couleurs en FUV	168
6.1	$r_0(z)$ pour la population globale, la population type 1+2 et la population type 3+4 dans le VVDS-02h	200
6.2	$r_0(z)$ des quatre types de populations définies dans le VVDS-02h	201
6.3	$r_0(z)$ des populations rouges et bleues définies dans le VVDS-02h	202
6.4	$r_0(z)$ des populations sélectionnées en NUV et FUV dans le VVDS-02h . .	203

Résumé

Les grands sondages récents de l'Univers ont mis en évidence la présence de structures dans la distribution de la matière, sous forme de filaments et de vides. Pour étudier l'évolution de la distribution spatiale des galaxies, il est nécessaire de connaître leur position précise dans l'espace à trois dimensions. Cette thèse s'est déroulée dans le cadre du sondage profond *VIMOS-VLT Deep Survey* (VVDS). Son but est de mesurer quelques 100000 *redshifts* pour étudier la formation et l'évolution des galaxies et des structures à grande échelle de l'Univers jusqu'à $z \sim 5$. Après avoir fait un état des lieux de la connaissance que nous avons de la distribution des galaxies, puis introduit le sondage VVDS, je présente la mesure et l'évolution de la fonction de corrélation spatiale à deux points à partir des données de la première époque du VVDS, le plus large échantillon (10759 spectres) jamais acquis à $I_{AB} = 24$. J'ai développé un ensemble de programmes mis à la disposition du consortium VVDS pour mesurer facilement la longueur de corrélation spatiale des galaxies dans un intervalle en *redshift* donné, avec ses erreurs associées, en corrigeant les effets de la stratégie d'observation du VVDS. Cet outil a permis de mesurer l'évolution de la fonction de corrélation spatiale de la population globale des galaxies jusqu'à $z=2$. J'ai prolongé cette étude en divisant l'échantillon de galaxies par type spectral et par couleur. Enfin, en combinant les données de GALEX avec celles du VVDS, j'ai pu mesurer l'agrégation de galaxies sélectionnées en ultraviolet jusqu'à $z \sim 1$. C'est la première fois que de telles mesures sont réalisées sur une si longue plage de temps cosmique. Les résultats présentés dans cette thèse font ainsi office de nouvelles références pour les travaux futurs.

Mot-clés : Astronomie - Astrophysique - Cosmologie - Sondages Profond - Spectroscopie - Simulations Numériques - Grandes structures - Galaxies - Univers

Abstract

The recent surveys of the Universe highlighted the presence of structures in the matter distribution, such as filaments and voids. To study the evolution of the galaxy spatial distribution, it is necessary to know their accurate position in a three dimensional space. This thesis took place within the framework of the *VIMOS-VLT Deep Survey* (VVDS). Its goal is to measure some 100000 redshifts to study the formation and evolution of the galaxies and large scale structures of the Universe up to $z \sim 5$. After having made a review of our knowledge of the galaxies distribution, then introduced the VVDS, I present the measurement and the evolution of the real-space two-point correlation function from the first epoch data of the VVDS, the largest sample of 10759 spectra ever acquired up to $I_{AB} = 24$. I developed a whole set of programs made available to the VVDS consortium to easily measure the clustering length of galaxies in a given redshift range, with its associated errors, correcting the effects of the VVDS observing strategy. This tool enabled to measure the evolution of the real space correlation function of the global galaxies population up to $z=2$. I then extended this study dividing the full galaxies sample by spectral type and color. Finally, combining the GALEX data to the VVDS has allowed me to measure the clustering of an ultraviolet-selected sample of galaxies up to $z \sim 1$. This is the first time that such measurements are carried out on such a so long range of cosmic time. The results presented in this thesis are thus establishing a new reference in the field.

Keywords : Astronomy - Astrophysics - Cosmology - Deep Surveys - Spectroscopy - Numerical Simulations - Large Structures - Galaxies - Universe

Laboratoire d'Astrophysique de Marseille
Traverse du Siphon - Les Trois Lucs - BP8
13376 Marseille Cedex 12 - France
